

DANIEL STEFFEN

**INTERVALO “BOOTSTRAP” PARA PREVISÕES DE SÉRIES
TEMPORAIS OBTIDAS PELO MÉTODO THETA**

Dissertação apresentada como requisito parcial à obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Métodos Numéricos em Engenharia, do Departamento de Construção Civil e do Departamento de Matemática da UFPR, na Área de Concentração em Programação Matemática e na Linha de Pesquisa em Métodos Estatísticos Aplicados à Engenharia sob a orientação do Prof. Dr. Anselmo Chaves Neto.

Curitiba

2010

TERMO DE APROVAÇÃO

DANIEL STEFFEN

INTERVALO “BOOTSTRAP” PARA PREVISÕES DE SÉRIES TEMPORAIS OBTIDAS PELO MÉTODO THETA

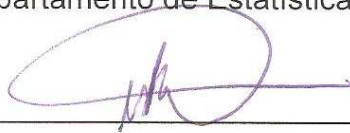
Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no curso de Pós-Graduação em Métodos Numéricos em Engenharia da Universidade Federal do Paraná, pela comissão formada pelos professores:

Orientador:



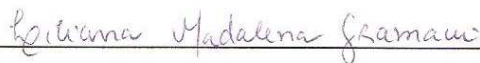
Prof. ANSELMO CHAVES NETO, D.Sc.

Departamento de Estatística da UFPR



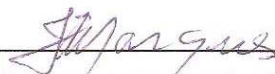
Prof. INÁCIO ANDRUSKI GUIMARÃES, D.Sc.

Departamento de Matemática da UTFPR



Profª. LILIANA MADALENA GRAMANI, Ph.D.

PPGMNE da UFPR



Prof. JAIR MENDES MARQUES, D.Sc.

PPGMNE da UFPR

Curitiba, 09 de julho de 2010.

AGRADECIMENTOS

Agradeço em primeiro lugar a Deus pela força que me deu durante o percurso de mais uma etapa vencida da minha vida.

Ao professor Dr. Anselmo Chaves Neto, pela orientação e disposição para com este trabalho.

Ao professor Dr. Jair Mendes Marques, pela ajuda prestada na programação em linguagem Matlab.

Aos meus pais Ernani e Anita Steffen pelo apoio irrestrito na minha trajetória, propiciando as condições para a conclusão deste curso.

Aos professores do PPGMNE da Universidade Federal do Paraná pelo esforço e dedicação no decorrer das disciplinas.

Aos colegas do mestrado, em especial a Maria Ivete Basniak e Tereza Rachel Mafioleti pela ajuda prestada no decorrer do curso.

Agradeço a todos que direta ou indiretamente contribuíram para a conclusão deste trabalho.

RESUMO

Este trabalho tem a finalidade de desenvolver um programa computacional em linguagem Matlab versão 7.1, do método de previsão de séries temporais chamado Theta. O método é relativamente recente e foi desenvolvido por Assimakopoulos e Nikolopoulos (2000). O Método Theta abordado neste trabalho mostrou-se muito eficiente na competição M3 de Makridakis, onde foram testadas 3003 séries temporais. Ele está baseado no conceito da modificação das curvaturas locais das séries temporais obtida por um coeficiente *theta* (Θ). Em sua abordagem mais simples a série temporal é decomposta em duas linhas *theta* $L(\Theta)$ representando os termos de longo prazo e de curto prazo. A previsão é feita combinando-se as previsões obtidas através do ajuste das linhas *theta* obtidas com a decomposição. Neste trabalho a série temporal é decomposta em até três linhas *theta*. As previsões obtidas já foram comparadas com métodos tradicionais e tiveram um bom desempenho. Mas, como desvantagem pode ser citada a falta de intervalos de confiança para as previsões. Neste trabalho é aplicada a técnica de reamostragem por computação intensiva conhecida como “Bootstrap” para estimar os intervalos de confiança para as previsões pontuais obtidas por esse método de previsão.

PALAVRAS-CHAVE: Método Theta, “Bootstrap”, Séries Temporais.

ABSTRACT

This work has porpose to develop a computer program in Matlab version 7.1, the method of time series forecasting called Theta. The method is relatively recent and was developed by Assimakopoulos and Nikolopoulos (2000). The Theta method presented in this work was very efficient in M3 Makridakis competition, where they were tested 3003 time series. It is based on the concept of modifying the local curvature of the time series obtained by a coefficient θ . In its simplest approach the time series is decomposed into two lines $L(\theta)$ representing terms of long-term and short term. The forecast is made by combining the forecasts obtained by fitting theta lines obtained with the decomposition. In this study the time series is decomposed into three lines theta. The forecasts obtained have been compared with traditional methods and had a good performance. But the disadvantage can be cited a lack of confidence intervals for the forecasts. This work applies the technique of computer-intensive resampling, known as the "bootstrap" to estimate confidence intervals for the point forecasts obtained by this method.

KEYWORDS: Theta method, "Bootstrap", Time Series.

LISTA DE FIGURAS

1 - ESTRUTURA DE FILTRO LINEAR.....	33
2 - FLUXOGRAMA DOS ESTÁGIOS DA METODOLOGIA BOX & JENKINS..	39
3 - ALGORITMO DA DISTRIBUIÇÃO “BOOTSTRAP” DE $T_n(x,F)$	51
4 - ALGORITMO DA DISTRIBUIÇÃO “BOOTSTRAP” DE $\hat{\beta}$	56
5 - ALGORITMO DA DISTRIBUIÇÃO “BOOTSTRAP” DE $\hat{Y}(h)$	63
6 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT21 PARA $\hat{Y}_T^*(1)$	72
7 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT22 PARA $\hat{Y}_T^*(1)$	76
8 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT23 PARA $\hat{Y}_T^*(1)$	80
9 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT31 PARA $\hat{Y}_T^*(1)$	85
10 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT31 PARA $\hat{Y}_T^*(1)$	91

LISTA DE QUADROS

1 - PROPRIEDADES DA FAC E FACP PARA MODELOS ARIMA.....	40
--	----

LISTA DE GRÁFICOS

1 - COTAÇÕES DIÁRIAS DO BOVESPA (JUL 2007/JUL 2009).....	17
2 - PADRÃO HORIZONTAL EM SÉRIES TEMPORAIS.....	21
3 - PADRÃO DE TENDÊNCIA EM SÉRIES TEMPORAIS.....	21
4 - PADRÃO CÍCLICO EM SÉRIES TEMPORAIS.....	22
5 - PADRÃO SAZONAL EM SÉRIES TEMPORAIS.....	22
6 - SÉRIE DO ÍNDICE BOVESPA E PREVISÃO THETA PELA DECOMPOSIÇÃO EM DUAS $L(\theta)$	45
7 - REPRESENTAÇÃO GRÁFICA DA SÉRIE TEMPORAL.....	66
8 - SÉRIE TEMPORAL, $L(\theta=0)^{RL}$, $L(\theta=2)^{AES}$ E PT21.....	70
9 - SÉRIE TEMPORAL, $L(\theta=0)^{RL}$, $L(\theta=2)^{AELB}$ E PT22.....	74
10 - SÉRIE TEMPORAL, $L(\theta=0)^{RL}$, $L(\theta=2)^{AEH}$ E PT23.....	78
11 - SÉRIE TEMPORAL, $L(\theta=0)^{RL}$, $L(\theta=0.5)^{AES}$, $L(\theta=2)^{AELB}$ E PT31.....	84
12 - SÉRIE TEMPORAL, $L(\theta=0)^{RL}$, $L(\theta=1.6)^{AELB}$, $L(\theta=2)^{AES}$ E PT32.....	89

LISTA DE TABELAS

1 - SÉRIE TEMPORAL GERADA POR SIMULADOR.....	66
2 - DECOMPOSIÇÃO DA SÉRIE EM DUAS LINHAS THETA.....	68
3 - EXTRAPOLAÇÃO DAS $L(\Theta)$ POR TENDÊNCIA LINEAR E ALISAMENTO EXPONENCIAL SIMPLES.....	69
4 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM).....	70
5 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO E O INTERVALO DE CONFIANÇA (NÍVEL DE SIGNIFICÂNCIA DE 5% E 1000 REPLICAÇÕES)...	71
6 - EXTRAPOLAÇÃO DAS $L(\Theta)$ POR TENDÊNCIA LINEAR E ALISAMENTO EXPONENCIAL LINEAR DE BROWN (AELB).....	73
7 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM).....	74
8 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO E O INTERVALO DE CONFIANÇA (NÍVEL DE SIGNIFICÂNCIA DE 5% E 1000 REPLICAÇÕES)...	75
9 - EXTRAPOLAÇÃO DAS $L(\Theta)$ POR TENDÊNCIA LINEAR E ALISAMENTO EXPONENCIAL DE HOLT (AEH).....	77
10 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM).....	78
11 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO E O INTERVALO DE CONFIANÇA (NÍVEL DE SIGNIFICÂNCIA DE 5% E 1000 REPLICAÇÕES)...	79
12 - DECOMPOSIÇÃO DA SÉRIE TEMPORAL EM TRÊS LINHAS THETA E COMBINAÇÃO.....	81

13 - EXTRAPOLAÇÃO DAS $L(\theta)$ POR TENDÊNCIA LINEAR, ALISAMENTO EXPONENCIAL SIMPLES (AES) E ALISAMENTO EXPONENCIAL LINEAR DE BROWN.....	82
14 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM).....	83
15 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO E O INTERVALO DE CONFIANÇA (NÍVEL DE SIGNIFICÂNCIA DE 5% E 1000 REPLICAÇÕES)...	84
16 - DECOMPOSIÇÃO DA SÉRIE TEMPORAL EM TRÊS LINHAS THETA E COMBINAÇÃO.....	87
17 - EXTRAPOLAÇÃO DAS $L(\theta)$ POR TENDÊNCIA LINEAR, ALISAMENTO EXPONENCIAL LINEAR DE BROWN (AELB) E ALISAMENTO EXPONENCIAL SIMPLES (AES)	88
18 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM).....	89
19 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO E O INTERVALO DE CONFIANÇA (NÍVEL DE SIGNIFICÂNCIA DE 5% E 1000 REPLICAÇÕES)...	90
20 - PREVISÕES E MAPE's OBTIDOS PELO MÉTODO THETA DECOMPOSTO EM DUAS LINHAS.....	91
21 - PREVISÕES E MAPE's OBTIDOS PELO MÉTODO THETA DECOMPOSTO EM TRÊS LINHAS E PREVISÕES E MAPE OBTIDO POR OUTRO MÉTODO DE PREVISÃO.....	92

SUMÁRIO

RESUMO.....	4
ABSTRACT	5
LISTA DE FIGURAS	6
LISTA DE QUADROS.....	6
LISTA DE GRÁFICOS.....	7
LISTA DE TABELAS.....	8
1 INTRODUÇÃO.....	13
1.1 PROBLEMA.....	14
1.2 OBJETIVOS	14
1.3 JUSTIFICATIVA.....	15
1.4 ESTRUTURA DA DISSERTAÇÃO.....	15
2 REVISÃO DE LITERATURA.....	16
2.1 TÉCNICAS DE PREVISÃO.....	16
2.2 SÉRIES TEMPORAIS.....	17
2.2.1 Definição.....	18
2.2.2 Objetivos.....	20
2.2.3 Padrões em Séries Temporais.....	20
2.3 MODELOS DE PREVISÃO QUANTITATIVOS.....	22
2.3.1 Tendência Linear.....	22
2.3.2 Modelos de Médias Móveis Simples (MMS).....	24
2.3.3 Alisamento Exponencial.....	26
2.3.3.1 Alisamento Exponencial Simples (AES).....	26
2.3.3.2 Alisamento Exponencial Linear de Brown (AELB).....	28
2.3.3.3 Alisamento Exponencial de Holt (AEH).....	30
2.4 METODOLOGIA DE BOX & JENKINS.....	32
2.4.1 Processos Estacionários.....	32
2.4.2 Modelo de Filtro Linear.....	33
2.4.3 Modelos Auto-Regressivos (AR).....	34
2.4.4 Modelos de Médias Móveis (MA).....	36
2.4.5 Modelos Auto-Regressivos e de Médias Móveis (ARMA).....	37

2.4.6 Modelos Auto-Regressivos Integrados de Médias Móveis (ARIMA).....	38
2.4.7 Identificação dos Modelos.....	40
2.5 MÉTODO THETA.....	41
2.5.1 Introdução.....	41
2.5.2 Considerações sobre o Método Theta.....	42
2.5.3 O Método Theta.....	42
2.5.4 Cálculo das Linhas Theta.....	45
2.6 TÉCNICA COMPUTACIONALMENTE INTENSIVA “BOOTSTRAP”	46
2.6.1 Introdução.....	46
2.6.2 Método Jackknife.....	47
2.6.3 Obtenção da Amostra Jackknife.....	48
2.6.4 Método “Bootstrap”	48
2.6.5 Obtenção da Amostra “Bootstrap”	50
2.6.6 Intervalos de Confiança “Bootstrap”	52
2.6.6.1 Intervalo t-student.....	53
2.6.6.2 Intervalo “Bootstrap” Residual.....	53
3 MATERIAL E MÉTODO.....	58
3.1 MATERIAL.....	58
3.1.1 Geração da Série Temporal.....	58
3.2 MÉTODO.....	59
3.2.1 Adaptação dos modelos e Implementação Computacional.....	59
3.2.1.1 Método de Previsão Theta.....	60
3.2.1.2 Previsões Theta.....	60
3.2.1.3 Intervalos de Confiança “Bootstrap”	60
4 RESULTADOS E DISCUSSÕES.....	65
4.1 PREVISÃO E INTERVALO “BOOTSTRAP” PARA A SÉRIE TEMPORAL.....	65
4.1.1 Descrição da Série Temporal.....	65
4.2 RESULTADOS PARA A SÉRIE DECOMPOSTA EM DUAS $L(\Theta)$	67
4.2.1 PT21 - Tendência Linear e Alisamento Exponencial Simples (AES).....	68
4.2.1.1 Previsões e Intervalo “Bootstrap” para PT21.....	69
4.2.2 PT22 - Tendência Linear e Alisamento Exponencial Linear de Brown (AELB).....	72

4.2.2.1 Previsões e Intervalo “Bootstrap” para PT22.....	73
4.2.3 PT23 - Tendência Linear e Alisamento Exponencial de Holt (AEH).....	76
4.2.3.1 Previsões e Intervalo “Bootstrap” para PT23.....	77
4.3 RESULTADOS PARA A SÉRIE DECOMPOSTA EM TRÊS $L(\Theta)$	80
4.3.1 PT31 - Tendência Linear, Alisamento Exponencial Simples (AES) e Alisamento Exponencial Linear de Brown (AELB).....	80
4.3.1.1 Previsões e Intervalo “Bootstrap” para PT31.....	83
4.3.2 PT32 - Tendência Linear, Alisamento Exponencial Linear de Brown (AELB) e Alisamento Exponencial Simples (AES).....	86
4.3.2.1 Previsões e Intervalo “Bootstrap” para PT32.....	88
4.4 COMPARAÇÃO ENTRE AS PREVISÕES OBTIDAS PELO MÉTODO THETA COM OUTROS MÉTODOS DE PREVISÃO.....	91
5 CONCLUSÃO E SUGESTÕES PARA TRABALHOS FUTUROS.....	93
REFERÊNCIAS.....	94
APÊNDICES.....	97
APÊNDICE A - PROGRAMA PARA A PREVISÃO DE SÉRIE TEMPORAL ATRAVÉS DO MÉTODO THETA UTILIZANDO O SOFTWARE MATLAB.....	97
APÊNDICE B - PROGRAMA PARA A PREVISÃO DE SÉRIE TEMPORAL ATRAVÉS DO MÉTODO DE ALISAMENTO EXPONENCIAL SIMPLES UTILIZANDO O SOFTWARE MATLAB.....	109
APÊNDICE C - PROGRAMA PARA A PREVISÃO DE SÉRIE TEMPORAL ATRAVÉS DO MÉTODO DE ALISAMENTO EXPONENCIAL DE BROWN UTILIZANDO O SOFTWARE MATLAB.....	110
APÊNDICE D - PROGRAMA PARA A PREVISÃO DE SÉRIE TEMPORAL ATRAVÉS DO MÉTODO DE ALISAMENTO EXPONENCIAL DE HOLT UTILIZANDO O SOFTWARE MATLAB.....	112
APÊNDICE E - PROGRAMA PARA A OTIMIZAÇÃO DO COEFICIENTE THETA UTILIZANDO O SOFTWARE MATLAB.....	114
APÊNDICE F - PROGRAMA GERADOR DE AMOSTRAS POR REAMOSTRAGEM COM REPOSIÇÃO.....	116

1 INTRODUÇÃO

O mundo vem passando por grandes mudanças sociais nas últimas décadas, mudanças que se tornam cada vez mais rápidas e imprevisíveis. Devido a essas mudanças, principalmente no que diz respeito a estratégias governamentais, finanças, marketing, produção agrícola, empresas de serviços, entre outras, pode-se avaliar a importância que os modelos de previsão tem na tomada de decisões. Essas previsões, tal qual são feitas pelas empresas para estimar a demanda de seus produtos e, conseqüentemente, planejar a programação da produção, de compras e de outras atividades inclusive identificar quando e onde concentrar os esforços de marketing (LIEBEL, 2004).

As técnicas de previsão podem ser divididas em duas categorias (ELSAYED e BOUCHER, 1994): as técnicas quantitativas e as técnicas qualitativas.

Neste trabalho é abordada uma técnica de previsão de séries temporais na qual as previsões são feitas a partir de sequências de valores passados, ou seja, das observações da série Z_t usando os valores $[Z_{t-1}, Z_{t-2}, Z_{t-3} \dots, Z_1]$. Os objetivos das técnicas de previsão de séries temporais são:

- Fazer previsão de valores futuros da série;
- Descrever apenas o comportamento da série (verificação de tendência e sazonalidade);
- Identificar o mecanismo gerador da série (o processo estocástico que gerou a série).

É bem conhecido que em séries temporais pode-se trabalhar com várias metodologias, e o desenvolvimento dos métodos através dos tempos pode ser esquematizado como:

- Métodos da era pré-Box e Jenkins (décadas de 30 e 40), onde se estudava e aplicava o método composto pelas componentes da tendência T_t , sazonalidade S_t e aleatoriedade A_t , ou seja, a série Z_t é uma função dessas três componentes, ou melhor, $Z_t = f(T_t, S_t, A_t)$.
- Na década de 60 surgiram os Métodos Automáticos ou “caixa-preta”, que são as técnicas de amortecimento exponencial.

- No início da década de 1970 surgiu o Método Box & Jenkins, quando G.E.P. Box & G. Jenkins publicaram os seus trabalhos que deram origem aos Modelos Box e Jenkins (1970).
- Na década de 1980 a inferência Bayesiana passou a ser aplicada em séries temporais e, mais recentemente surgiram outras técnicas. Uma que tem alcançado bastante sucesso é o método das Redes Neurais (computacionalmente intensivo) que passou a ser usado também em previsão.

1.1 PROBLEMA

Os modelos de previsão têm grande importância nos meios acadêmicos e sociais devido a sua vasta aplicabilidade nas diversas áreas do conhecimento científico, industrial, comercial e de serviços. Uma ilustração interessante é a construção de usina hidroelétrica onde o conhecimento da série temporal da vazão do rio que abastecerá a represa é fundamental ao desenvolvimento do projeto. Atualmente, as indústrias necessitam planejar em detalhes a sua produção e o estoque mantido a disposição das operações. Deste modo a aplicação de séries temporais é fundamental.

O Método Theta, abordado neste trabalho, possibilita decompor séries temporais em um conjunto de novas séries antes da aplicação de um modelo de previsão. No entanto, as previsões feitas pelo método Theta são pontuais e ainda não foi desenvolvido uma estimativa por intervalo para essas previsões. Esse é o problema fundamental desse método.

1.2 OBJETIVOS

O principal objetivo deste trabalho é aplicar a técnica de reamostragem, por computação intensiva, conhecida como “Bootstrap” para obter um intervalo de confiança para as previsões pontuais fornecidas pelo método Theta de previsão. E, como objetivo específico será desenvolvido um programa computacional experimental em ambiente Matlab versão 7.1 para ajustar o modelo de previsão e verificar a eficiência do mesmo.

1.3 JUSTIFICATIVA

Este trabalho se justifica pelo fato de que a aplicação de técnicas de previsão de séries temporais, comprovadamente eficientes e aliadas à informática vem a favorecer inúmeros setores da sociedade, pois é necessário ter informações sobre números futuros em todos os setores dos empreendimentos que merecem mais atenção, tanto em fatores de investimento como em realocação de pessoal e material.

A aplicação do Método Theta de previsão garante uma grande aplicabilidade em empresas de pequeno porte, pois as mesmas não possuem bases estatísticas para avaliar seus dados, ou seja, a utilização de métodos com eficácia comprovada permite que pequenas empresas utilizem modelos de previsão sem grandes custos e complicações, pois são fáceis de serem aplicados.

1.4 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação é composta por essa introdução, de uma revisão de literatura no capítulo 2, onde se aborda o método de previsão Theta, objeto de estudo desta dissertação, e o método de computação intensiva conhecido por “Bootstrap”. Já no capítulo 3 são descritos o material e o método aplicado no trabalho. No capítulo 4 têm-se os resultados e a sua discussão. Finalmente, tem-se no capítulo 5 a conclusão.

2 REVISÃO DE LITERATURA

2.1 TÉCNICAS DE PREVISÃO

Uma previsão nada mais é que uma estimativa quantitativa sobre eventos futuros baseados em informações de períodos passados e recentes. As técnicas de previsão podem ser divididas em duas categorias (ELSAYED & BOUCHER, 1994): técnicas qualitativas e quantitativas.

Técnicas qualitativas são usadas quando há pouca informação e, assim, é necessário consultar especialistas no assunto, e suas opiniões tendem a ser a palavra final da previsão.

Técnicas quantitativas são usadas quando existe um histórico de dados (venda, receita, custo, produção, demanda para dimensionamento de estoque, entre outros), e é possível aplicar diferentes métodos estatísticos com o objetivo de extrapolar esses dados e obter prováveis valores futuros.

As duas principais técnicas são as séries temporais e os modelos estruturais. Este último surgiu como uma alternativa aos modelos de Box & Jenkins (VICENTE, 1992). Em tais modelos, Y_t é decomposto em termos da tendência (μ_t), do ciclo (Ψ_t), da sazonalidade (γ_t) e de um componente irregular (ε_t), ou seja,

$$Y_t = \mu_t + \Psi_t + \gamma_t + \varepsilon_t \quad (1)$$

A equação (1) é conhecida como equação das medidas ou equação das observações. (BRESSAN, 2002)

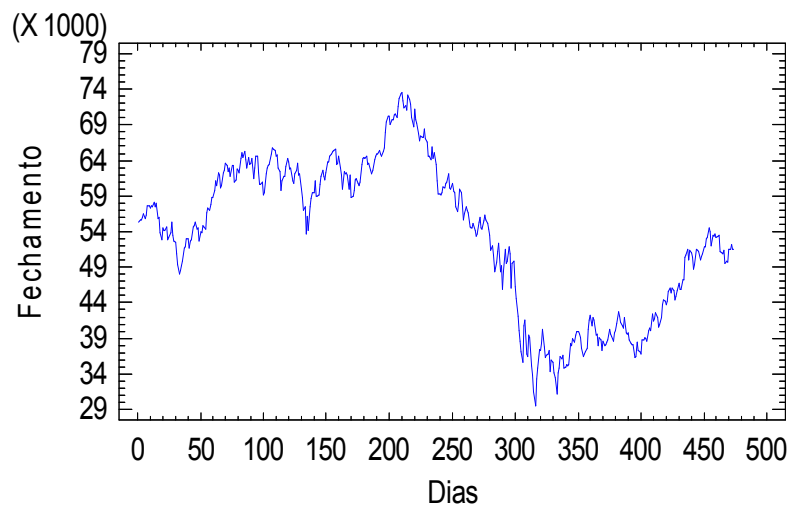
A análise de séries temporais compreende modelos com os quais se realizam previsões a partir de sequências de dados históricos, ou seja, são previsões feitas em função de realizações passadas. Nos modelos estruturais a análise é feita a partir das observações de variáveis emparelhadas (STEVENSON, 1981). Neste estudo aborda-se as séries temporais.

2.2 SÉRIES TEMPORAIS

O conceito de séries temporais está relacionado a um conjunto de observações de uma determinada variável feita em períodos sucessivos de tempo e ao longo de um determinado intervalo. É possível citar como exemplos as cotações diárias da taxa do dólar, as vendas mensais de um produto, a taxa de desemprego de um país, entre outros.

No ajuste de modelos de séries temporais o ponto de partida é fazer a representação gráfica da série, pois através do gráfico se pode fazer uma melhor análise e identificar as características que podem ser relevantes para o estudo da série em questão. Um exemplo bem conhecido de séries temporais são os índices diários da bolsa de valores de São Paulo (Índice Bovespa):

GRÁFICO 1 - COTAÇÕES DIÁRIAS IBOVESPA (JULHO 2007/JULHO 2009)



FONTE: <http://www.bolsapt.com/historico/%5EBVSP/25/de-01-07-2007-a-01-07-2009/>

Existem vários métodos para a previsão de séries temporais, a escolha de métodos mais sofisticados não necessariamente apresenta melhores resultados. Métodos mais simples podem render resultados muito satisfatórios dependendo de certas condições, além de permitir uma melhor compreensão dos resultados. Sendo assim, deve-se primeiro avaliar o tipo da série para então verificar a necessidade de aplicação de métodos mais complexos.

2.2.1 Definição

Uma série temporal, segundo CHAVES NETO (2004), é um conjunto de realizações geradas sequencialmente no tempo e que apresentam uma dependência serial (isto é, dependência entre instantes de tempo). A notação usada para denotar uma série temporal é $Z_1, Z_2, Z_3, \dots, Z_T$, que indica uma série de tamanho T . O instante T indica o último instante disponível.

A característica mais importante deste tipo de dados é que as observações vizinhas são dependentes e estamos interessados em analisar e modelar esta dependência. Enquanto em modelos de regressão, por exemplo, a ordem das observações é irrelevante para a análise, em séries temporais a ordem dos dados é crucial (EHLERS 2007).

A maior parte dos procedimentos estatísticos foi desenvolvida para analisar observações independentes, então o estudo de séries temporais requer o uso de técnicas específicas que consideram a correlação entre as realizações nos instantes de tempo. Dados de séries temporais surgem em vários campos do conhecimento e existem algumas características particulares a esse tipo de dados:

- Observações correlacionadas são mais difíceis de analisar e requerem técnicas específicas.
- Precisa-se levar em conta a ordem temporal das observações.
- Fatores complicadores como presença de tendências e variação sazonal ou cíclica, podem ser difíceis de estimar ou remover.
- A seleção de modelos pode ser bastante complicada, e as ferramentas podem ser de difícil interpretação.
- É mais difícil de lidar com observações perdidas e dados divergentes, devido à natureza seqüencial.

Quanto aos dados existem dois tipos de séries, são elas: as discretas e as contínuas. Uma série temporal é dita ser contínua quando as observações são feitas continuamente no tempo. Um exemplo clássico de uma série contínua é a vazão de um rio ou, então, a altura das marés. Essas séries contínuas podem ser discretizadas tomando-se valores a instantes de tempo fixados. Definindo o conjunto

de índices de períodos como $T = \{t : t_1 < t < t_2\}$ a série temporal pode ser denotada por $\{X(t) : t \in T\}$. Uma série temporal é dita ser discreta quando as observações são feitas em tempos específicos, geralmente equiespaçados. Definindo o conjunto $T = \{t_1, \dots, t_n\}$ a série temporal será denotada por $\{X_t : t \in T\}$.

Uma série temporal também pode ser multivariada. Se k variáveis são observadas a cada instante de tempo (valores discretizados) denota-se por $\{X_{1t}, \dots, X_{kt}, t \in T\}$. Neste caso várias séries correlacionadas devem ser analisadas conjuntamente, ou seja, em cada instante de tempo tem-se um vetor de observações.

Segundo CHAVES NETO (2004), considerava-se a série temporal $\{Z_t, t = 1, 2, 3, \dots, n\}$ como composta por 4 componentes distintas: T_t (Tendência), S_t (Sazonalidade), C_t (Ciclo), a_t (ruído aleatório). As formas de decomposição dessas componentes são:

$$\text{Modelo aditivo: } Z_t = T_t + S_t + C_t + a_t$$

$$\text{Modelo mutiplicativo: } Z_t = T_t \cdot S_t \cdot C_t \cdot a_t$$

$$\text{Modelo misto: } Z_t = T_t \cdot S_t \cdot C_t + a_t$$

H. O. Wold em 1938 mostrou que qualquer série temporal Z_t discreta poderia ser representada por modelos AR (auto-regressivos) e MA (médias móveis), (WOLD, 1938). Porém, só foi possível a implementação destes métodos na década de 60 com o advento dos computadores de 2ª geração (transistor).

Nas décadas de 50 e 60 foram desenvolvidos os métodos automáticos (Brown e Winters, entre outros), os quais podem ser programados em computador e que não requerem a intervenção de analistas. Os métodos de alisamento exponencial constituem a formulação mais popular desta categoria.

2.2.2 Objetivos

Segundo MORETTIN & TOLOI (1981) os objetivos de se analisar uma série temporal são os seguintes:

- 1) Descrição: propriedades da série como, por exemplo, o padrão de tendência, a existência de alterações estruturais, etc.
- 2) Explicação: construir modelos que permitam explicar o comportamento da série no período observado.
- 3) Controle de Processos: por exemplo, controle estatístico de qualidade.
- 4) Previsão: prever valores futuros com base em valores passados.

Para alcançar esses objetivos, necessita-se identificar o padrão gerador dos dados da seqüência observada. Em seguida, extrapola-se o padrão identificado e se obtêm previsões de eventos futuros (STATSOFT, 1999).

2.2.3 Padrões nas Séries Temporais

KRAJEWSKI & RITZMAN (1998) apresentam cinco padrões básicos, observados em séries temporais:

(i) Linearidade: é a série temporal no qual ocorre flutuação dos dados em torno de uma média constante;

(ii) Tendência: ocorre quando a série temporal apresenta um acréscimo ou decréscimo em sua média ao longo do tempo;

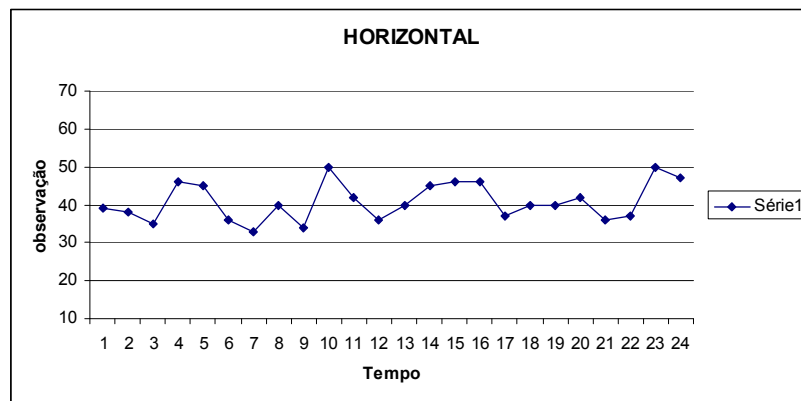
(iii) Sazonal: é um padrão de acréscimos e decréscimos que se repete em períodos determinados de tempo da série temporal, que podem ser em dias, semanas, meses ou anos;

(iv) Cíclicos: são padrões de acréscimos ou decréscimos graduais no valor da média de observações da série temporal em períodos mais longos de tempo;

(v) Randômico: ocorre quando não se pode prever o comportamento da série.

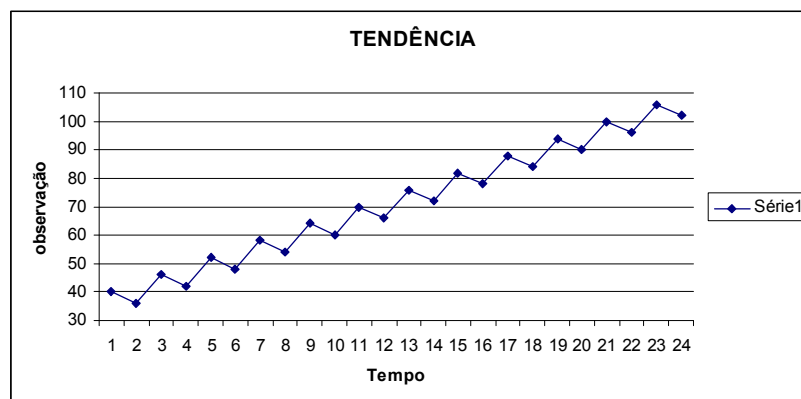
Os quatro primeiros padrões apresentados acima, podem ser encontrados em diversas combinações que acabam resultando num padrão identificável e previsível em séries temporais. Já o padrão randômico ocorre ao acaso e não pode ser previsto. Quando o padrão randômico ocorrer em uma série com padrões definidos e previsíveis, as observações podem ser consideradas especiais, necessitando uma análise mais aprofundada para entender a sua ocorrência. Adiante apresenta-se vários padrões de realizações de séries temporais.

GRÁFICO 2 - PADRÃO HORIZONTAL DE SÉRIES TEMPORAIS



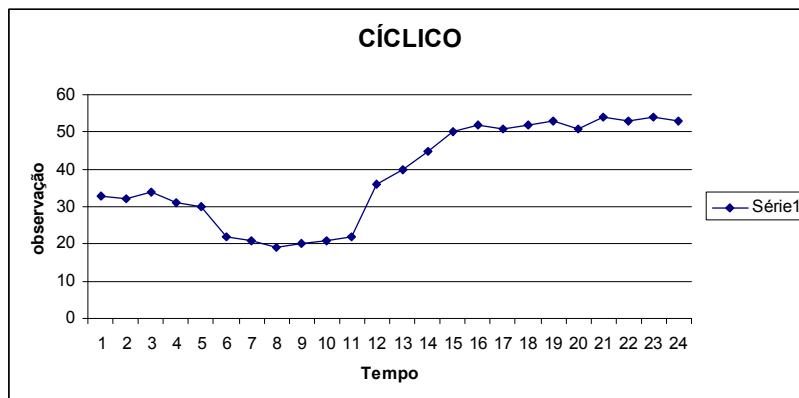
FONTE: O Autor (2010)

GRÁFICO 3 - PADRÃO DE TENDÊNCIA DE SÉRIES TEMPORAIS



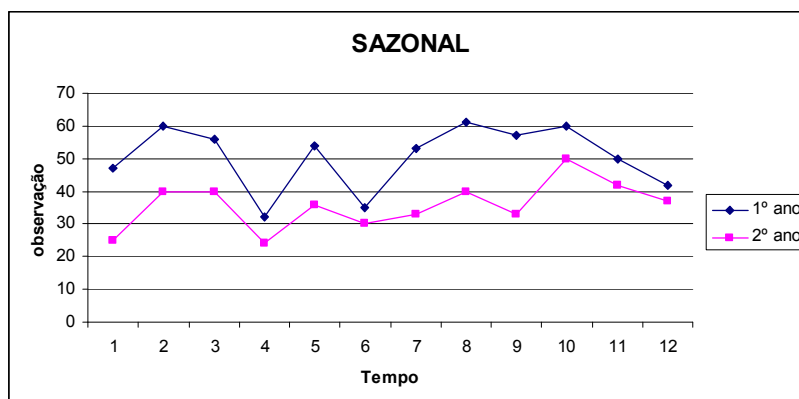
FONTE: O Autor (2010)

GRÁFICO 4 - PADRÃO CÍCLICO DE SÉRIES TEMPORAIS



FONTE: O Autor (2010)

GRÁFICO 5 - PADRÃO SAZONAL DE SÉRIES TEMPORAIS



FONTE: O Autor (2010)

2.3 MODELOS DE PREVISÃO QUANTITATIVOS

2.3.1 Tendência Linear

Tendência Linear é uma das técnicas mais difundidas e utilizadas na atualidade (PELEGRINI, 2000). Segundo MAKRIDAKIS (*et al*, 1998) a equação genérica de regressão linear simples, adaptada ao contexto de previsão, é dada por:

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

onde, Y_i representa a $i^{\text{ésima}}$ observação de Y , X_i representa a $i^{\text{ésima}}$ observação de X , α e β são constantes de suavização (desconhecidas) e ε_i é uma variável randômica que segue a distribuição normal com média zero e variância σ_ε^2 . A relação deste tipo de regressão é representada por uma reta, logo as previsões são pontos dessa reta. Para que sejam feitas extrapolações é necessário que a correlação entre X e Y seja suficientemente forte, garantindo um valor do coeficiente de correlação R^2 próximo de 1 (LIEBEL, 2004).

O coeficiente de correlação mede o ajuste do modelo proposto aos dados amostrais, ou seja, quanto o modelo explica da variação dos dados presente na parte estocástica do modelo. Valores próximos a zero indicam que o modelo não é adequado à série.

Em regressão linear simples o objetivo é estimar os parâmetros α e β , ou seja, determinar as estimativas dos parâmetros, α e β , que minimizam o erro. O erro ε_i é a diferença entre o valor da variável dependente Y_i observada e sua estimativa \hat{Y}_i fornecida pelo modelo. O método dos mínimos quadrados é o mais utilizado para estimar os parâmetros dessa regressão (STEVENSON, 1981). Supondo o modelo de regressão escrito na forma matricial:

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon},$$

deve-se estimar o vetor de parâmetros $\underline{\beta}$ para se modelar os dados. O melhor estimador do vetor de parâmetros $\underline{\beta}$ é chamado de BLUE (melhor estimador linear não viciado) e tem por expressão:

$$\hat{\underline{\beta}} = (X'X)^{-1}X'\underline{Y}. \quad (2)$$

O caso geral descrito na expressão (2) é conhecido como regressão linear múltipla, que é simplesmente a extensão do modelo de regressão linear simples para um número maior de variáveis.

2.3.2 Modelos de Médias Móveis Simples (MMS)

Considerando a série temporal Z_1, \dots, Z_t , localmente constante (sem tendência e sazonalidade), pode-se expressá-la em função de seu nível mais um ruído aleatório (ruído branco), ou seja:

$$Z_t = \mu_t + a_t, \quad t = 1, \dots, N,$$

onde $E(a_t) = 0$, $V(a_t) = \sigma_a^2$ e μ_t é um parâmetro desconhecido, que pode variar lentamente com o tempo e representa o nível da série. Usando-se o procedimento de média móvel o que se faz é calcular a média aritmética das k observações mais recentes, ou seja:

$$M_t = \frac{Z_t + Z_{t-1} + Z_{t-2} + \dots + Z_{t-k+1}}{k}$$

ou

$$M_t = M_{t-1} + \frac{Z_t - Z_{t-k}}{k}. \quad (3)$$

Assim, a equação (3) é uma estimativa do nível μ_t que não pondera observações mais antigas e sim prioriza as mais recentes. Devido a esse fator atribui-se o nome de média móvel, pois a cada período, a observação mais antiga é substituída pela mais recente.

A previsão de todos os valores futuros é dada pela última média móvel, isto é:

$$\hat{Z}_t(h) = \hat{Z}_{t-1}(h+1) + \frac{Z_t - Z_{t-k}}{k}, \quad (4)$$

para todo horizonte $h > 0$. Na verdade a expressão (4) consiste num algoritmo de atualização da previsão, devido que a cada instante, ou melhor, a cada nova observação é feita uma correção da estimativa anterior de Z_{t+h} . A média e a variância da previsão são dados por:

$$E[\hat{Z}_t(h)] = \mu \quad \text{e} \quad V[\hat{Z}_t(h)] = \frac{\sigma_a^2}{k^2}.$$

E, assumindo Gaussianidade para os resíduos, $a_t \sim N(0, \sigma_a^2)$, tem-se que $\hat{Z}_t(h) \sim N(\mu, \frac{\sigma_a^2}{k^2})$ (CHAVES NETO, 2004). Um dos problemas deste método é a escolha de k , quanto maior o valor de k mais suave é a previsão. Se k é pequeno a previsão oscila muito. Uma característica importante do método de médias móveis é: todas as observações têm o mesmo peso $\frac{1}{k}$, na prática é razoável supor que as observações mais recentes sejam mais relevantes para os próximos valores da série, e portanto devem receber um peso maior que as observações mais antigas. O melhor tamanho de k é aquele que minimiza o erro quadrático médio de previsão, ou seja, o EQM.

Segundo MORETTIN & TOLOI (2006) dentre as principais vantagens deste método podemos citar a simplicidade e facilidade de aplicação. Esse método é recomendado quando se tem um número pequeno de observações, ou seja, uma série curta. E possui uma boa característica, que é a grande flexibilidade devido à possibilidade de escolha de k de acordo com o padrão da série.

MORETTIN & TOLOI (2006) também ressaltam algumas desvantagens que este método apresenta. Ele é utilizado somente para prever séries estacionárias, caso contrário a precisão obtida será muito pequena, pois os pesos atribuídos às k observações são todos iguais e nenhum peso é dado às observações anteriores. Outra dificuldade é a necessidade de armazenar pelo menos $(k - 1)$ observações.

Atualmente este método não é muito utilizado, pois o método de Alisamento Exponencial Simples possui todas as vantagens e ainda mais benefícios, que são descritos a seguir.

2.3.3 Alisamento Exponencial

A teoria fundamental do alisamento exponencial foi construída por BROWN e MEYER (1961) e nela estabeleceu-se que, em qualquer série temporal $\{Z_t, t = 0, 1, \dots, n\}$, deve existir no tempo t um único polinômio representando a série temporal:

$$Z_{t+m} = a_t + b_t m + \frac{c_t}{2} m^2 + \dots + \frac{g_t}{K} m^K, \quad (5)$$

onde Z_{t+m} é a previsão para o horizonte m estando no instante t e $a_t, b_t, c_t, \dots, g_t$ são os coeficientes da combinação linear da equação. Os coeficientes da equação (5) podem ser estimados como combinações lineares dos valores resultantes dos primeiros $k+1$ graus (simples, duplo, triplo, etc.) de alisamento aplicado para os valores Z_t .

Os métodos de alisamento são os mais utilizados pelas empresas devido a simplicidade na implementação e geração de bons resultados. Os modelos de Alisamento Exponencial podem ser de vários tipos: modelos de alisamento exponencial simples, modelos de alisamento exponencial Linear de Brown (AELB), modelos de Holt-Winters com e sem sazonalidade, dentre outros.

2.3.3.1 Alisamento Exponencial Simples (AES)

A idéia geral do alisamento (amortecimento) exponencial é parecida com a do método de médias móveis, mas os pesos das observações decrescem à medida que as observações estão mais longe do passado. Os pesos atribuídos aos elementos da série temporal decaem exponencialmente (razão do nome suavização exponencial), do mais recente para o mais antigo (PELLEGRINI & FOGLIATTO, 2002).

A maior dificuldade na aplicação é escolher a(s) constante(s) de alisamento, mas alguns Softwares já ajustam esses modelos automaticamente com constantes de amortecimento otimizadas (BARROS, 2006).

O método de alisamento exponencial simples (AES) se utiliza de médias móveis ponderadas exponencialmente (MAKRIDAKIS, *et al.*, 1998). O argumento

para o tratamento diferenciado das observações da série temporal é fundamentado na suposição de que as últimas observações contêm mais informações sobre o futuro e, portanto, são mais relevantes para a previsão.

MORETTIN & TOLOI (2006) especificam o método AES através da seguinte equação:

$$\hat{Z}_t = \alpha Z_t + (1 - \alpha) \hat{Z}_{t-1}, \quad t = 1, \dots, N \quad (6)$$

onde, \hat{Z}_t é o valor exponencialmente alisado (suavizado) e α é a constante de alisamento, $0 < \alpha < 1$. Pode-se demonstrar que a equação (6) se reduz a $\hat{Z}_{t+1} = \alpha e_t + \hat{Z}_t$, onde $e_t = Z_t - \hat{Z}_t$ é o erro de previsão a um passo. Assim, a nova previsão pode ser obtida da anterior adicionando-se um múltiplo do erro de previsão. A soma dos pesos atribuídos às observações converge para 1 quando $t \rightarrow \infty$:

$$\alpha[1 + (1 - \alpha) + (1 - \alpha)^2 + \dots] = \alpha + \alpha(1 - \alpha) + \alpha(1 - \alpha)^2 + \dots,$$

ocorrendo a soma de uma progressão geométrica com $q = (1 - \alpha)$ de onde se obtém,

$$S = \frac{\alpha}{1 - (1 - \alpha)} = 1, \quad 0 < \alpha < 1$$

A constante de ponderação α está limitada ao intervalo aberto 0 e 1. Quanto maior o seu valor maior será o peso dado às observações mais recentes (CORRÊA; GIANESI; CAON, 2000). Valores de α pequenos produzem previsões que dependem mais das observações passadas, pois o peso das mesmas será então maior. Enquanto valores de α próximos de 1 dependem das observações mais recentes, que são as observações que recebem o maior peso. De maneira semelhante à utilizada no método da Regressão Linear Simples, busca-se um valor de α que minimize o EQM.

Segundo CHAVES NETO (2004), para se prever valor futuro da série usando o (AES), onde usa-se o último valor amortecido exponencialmente, ou seja,

$$\hat{\mu}_t(h) = \hat{\mu}_t \quad \forall h,$$

ou

$$\hat{\mu}_t(h) = \alpha Z_T + (1 - \alpha) \hat{\mu}_{t-1}(h+1).$$

A esperança do estimador é $E[\hat{\mu}_t(h)] = \mu$ e a variância é $V[\hat{\mu}_t(h)] = \frac{\alpha}{2 - \alpha} \sigma_a^2$ e assumindo Gaussianidade tem-se o intervalo de confiança,

$$P[\hat{\mu}_t(h) - z_c \sqrt{\frac{\alpha}{2 - \alpha} \sigma_a^2} \leq \mu \leq \hat{\mu}_t(h) + z_c \sqrt{\frac{\alpha}{2 - \alpha} \sigma_a^2} = 1 - \alpha.$$

O método agrega duas qualidades essenciais em um contexto de previsão de demanda em empresas: é simples e requer poucos dados. Administrativamente isso implica em pouco gasto com treinamento de pessoal para sua utilização e com coleta de dados (LIEBEL, 2004).

2.3.3.2 Alisamento Exponencial Linear de Brown (AELB)

Quando uma série apresenta tendência o AES não é adequado, pois fornece previsões que subestimam ou superestimam constantemente os valores reais. Para evitar essas falhas utiliza-se o método de Alisamento Exponencial Linear de Brown.

De acordo com BARROS (2006), o Alisamento Exponencial Linear de Brown é um método semelhante ao anterior, indicado a séries que apresentam tendência. Nele, uma única constante de alisamento é empregada no processo de atualização de ambos os parâmetros m_1 e b_2 .

Este método consiste em calcular um segundo valor exponencialmente alisado ($\bar{\bar{Z}}_t$) pela equação

$$\bar{\bar{Z}}_t = \alpha \bar{Z}_t + (1 - \alpha) \bar{\bar{Z}}_{t-1}, \quad \tilde{\tilde{Z}}_1 = Z_1$$

onde

$\bar{\bar{Z}}_t$: valor duplamente alisado;

\bar{Z}_t : alisamento exponencial simples;

α : constante de alisamento entre 0 e 1;

Considerando-se que a tendência seja linear, o modelo poderá ser determinado por:

$$Z_t = m_1 + b_2 t + a_t, \quad t=1, \dots, N,$$

onde m_1 é o intercepto e b_2 é a declividade. A equação de previsão é:

$$\hat{Z}(h) = \hat{m}_{1,t} + \hat{b}_{2,t}(h),$$

onde:

$$\hat{m}_{1,t} = 2\tilde{\tilde{Z}}_t - \bar{\bar{Z}}_t,$$

$$\hat{b}_{2,t} = \frac{\alpha}{1 - \alpha} \left(\tilde{\tilde{Z}}_t - \bar{\bar{Z}}_t \right).$$

Em todos os métodos de alisamento é necessário escolher valores iniciais dos parâmetros de forma a inicializar as equações de atualização. Algumas escolhas possíveis são:

$$\bar{\bar{Z}}_1 = Z_1$$

$$\hat{m}_1(1) = Z_1$$

$$b_2(1) = \frac{(Z_2 - Z_1) + (Z_4 - Z_3)}{2}.$$

A previsão k passos à frente no instante t é:

$$\hat{Z}_{t+k} = \hat{m}_1(t) + \hat{b}_2(t) \cdot (k).$$

Segundo MORETTIN & TOLOI (1981), as vantagens do método são a grande flexibilidade de ajuste devido à constante de alisamento, facilidade de implementação e entendimento, sendo mais eficiente em termos computacionais do que as médias móveis. A desvantagem é que ele é direcionado para a tendência linear e é menos flexível que o modelo de Holt que usa duas constantes de alisamento.

2.3.3.3 Alisamento Exponencial de Holt (AEH)

Este método é parecido com o anterior, mas agora existem duas constantes de alisamento diferentes que alisam diretamente os valores da tendência.

Os valores do nível e da tendência da série, no instante t, serão estimados por:

$$\bar{Z}_t = AZ_t + (1 - A)(\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1 \text{ e } t = 2, \dots, N$$

$$\hat{T}_t = C(\bar{Z}_t - \bar{Z}_{t-1}) + (1 - C)\hat{T}_{t-1}, \quad 0 < C < 1 \text{ e } t = 2, \dots, N$$

respectivamente. A e C são denominados constantes de alisamento. As fórmulas \bar{Z}_t e \hat{T}_t como em todos os métodos de alisamento, modificam estimativas prévias quando uma nova observação é obtida (MORETTIN; TOLOI, 2006).

A previsão para a observação Z_{t+h} , no instante t é dada por:

$$\hat{Z}_t(h) = \tilde{Z}_t + h\hat{T}_t, \quad \forall h > 0$$

a previsão é feita adicionando-se ao valor \tilde{Z}_t a tendência multiplicada pelo número de passos à frente que se deseja prever (h). As equações utilizadas para atualização da previsão quando se tem uma nova observação Z_{t+1} são:

$$\tilde{Z}_{t+1} = AZ_{t+1} + (1-A)(\tilde{Z}_t + \hat{T}_t), \quad (7)$$

$$\hat{T}_{t+1} = C(\tilde{Z}_{t+1} - \tilde{Z}_t) + (1-C)\hat{T}_t, \quad (8)$$

a previsão para Z_{t+h} será:

$$\hat{Z}_{t+1}(h-1) = \tilde{Z}_{t+1} + (h-1)\hat{T}_{t+1}.$$

Para utilizar as equações (7) e (8) é necessário atribuir valores iniciais para \hat{T}_2 e \tilde{Z}_2 . A forma mais simples de estimar os valores de \hat{T}_2 e \tilde{Z}_2 é dada por:

$$\hat{T}_2 = Z_2 - Z_1 \text{ e } \tilde{Z}_2 = Z_2.$$

Para determinar as constantes de alisamento A e C deve-se utilizar um instrumento de comparação entre os valores previstos e realizados, como a média dos quadrados dos erros (MQE). Os valores ideais são aqueles que minimizam o MQE.

2.4 METODOLOGIA DE BOX & JENKINS

Os modelos de Box & Jenkins também conhecidos como modelos ARIMA, foram desenvolvidos nos anos 70, com base nos trabalhos de Yule que introduziu os modelos autoregressivos (AR) em 1926 e os modelos de médias móveis (MA) introduzido por Slutsky em 1937. A partir daí Box & Jenkins construíram toda uma técnica de identificação do modelo, estimação de parâmetros e verificação da validade do modelo. Estes modelos consistem em uma classe de modelos lineares que associam de forma conjunta parte auto-regressiva e parte médias móveis, ou seja, é faz-se uma integração dos processos (TRENTIN, 2002).

Contudo existem algumas restrições na aplicação da técnica, tais como o número de amostras que deve ser maior que 50 observações e a necessidade de intervenção de um analista no trato das informações; (CHAVES NETO, 1991).

2.4.1 Processos Estacionários

Um processo é dito estacionário quando tem média e variância constantes, independentemente da escolha da origem dos tempos (CHAVES NETO, 2004), ou seja, $\mu(t)=\mu$ e $V(t)=\sigma^2$.

A estacionariedade ou não-estacionariedade de uma série temporal pode ser verificada visualmente pela análise gráfica das observações da série. Quando as observações da série oscilam em torno da média, a série é dita estacionária. Trata-se de uma série denominada ruído aleatório composta de nível da série (c) e erro aleatório (a_t). MAKRIDAKIS *et al.* (1998) apresentam o modelo matemático que representa a série estacionária de ruído aleatório,

$$Z_t = c + a_t$$

A aplicação dos modelos de Box & Jenkins necessita em um primeiro momento transformar séries não-estacionárias em estacionárias, se necessário. Para isso, deve-se remover os padrões não necessários na série tomando-se um número finito de diferenças d (MORETTIN; TOLOI, 1981). A diferenciação é feita utilizando-se um operador de retardo B , tal que,

$$BZ_t = Z_{t-1},$$

então tem-se:

$$\nabla Z_t = Z_t - Z_{t-1} = (1 - B)Z_t,$$

portanto:

$$\nabla = 1 - B.$$

Se a ordem da diferenciação é $d = 2$, tem-se:

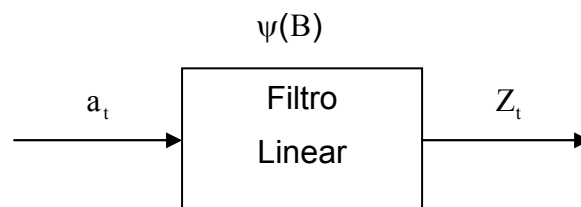
$$\omega_t = (Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2}) = Z_t - 2Z_{t-1} + Z_{t-2} = \nabla^2 Z_t$$

Na maioria dos casos a estacionariedade da série é obtida após uma ou duas diferenças (MORETTIN & TOLOI, 1981).

2.4.2 Modelo de Filtro Linear

Este modelo supõe que a série temporal seja gerada através de um filtro linear ou sistema linear, cuja entrada é ruído branco a_t (MORETTIN & TOLOI, 2006).

FIGURA 1 - ESTRUTURA DE FILTRO LINEAR



Este modelo consiste na soma das observações aleatórias a_t anteriores com atribuição de diferentes pesos a cada observação, como mostrado na equação:

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \psi(B)a_t ,$$

em que

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots \quad (9)$$

A equação (9) denominada função de transferência do filtro é responsável pela transformação de a_t em Z_t , e μ é um parâmetro determinando o nível da série (MORETTIN & TOLOI, 2006).

De acordo com a análise de LIEBEL (2004) a sequência dos pesos ψ pode ser finita ou infinita. Independente disso, a soma desses pesos será determinante para a análise da série. Se esta soma tiver um resultado menor que o infinito, então o filtro é estável e o processo Z_t é estacionário, se Z_t é estacionário, μ é a média ao redor da qual o processo varia.

2.4.3 Modelos Auto-regressivos (AR)

Os modelos auto-regressivos são um caso especial de filtro linear. A principal característica do modelo auto-regressivo é que em lugar das variáveis independentes o processo se utilizará dos valores prévios da série para estimação do modelo; (LIEBEL, 2004).

Considerando a série temporal Z , com observações $Z_t, Z_{t-1}, Z_{t-2}, \dots$, e seus desvios da média $\tilde{Z}_t, \tilde{Z}_{t-1}, \tilde{Z}_{t-2}, \dots$, dados por:

$$\tilde{Z}_t = Z_t - \mu$$

Então o chamado processo auto-regressivo de ordem p , simbolicamente definido por $AR(p)$ é estimado por:

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t ,$$

onde:

Z_t é o valor observado na série no instante t ;

δ representa o termo constante da série;

ϕ_i corresponde ao i -ésimo parâmetro auto-regressivo, $i = 1, 2, \dots, p$;

a_t é o erro (ruído) no instante t ;

Para a simplificação do modelo, utiliza-se um operador auto-regressivo de ordem p , dado por:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

A estacionariedade é garantida se todas as raízes da equação $\phi(B) = 0$, com base no polinômio apresentado pela equação acima forem maiores do que 1 em módulo.

Para os processos auto-regressivos de primeira ordem, AR(1), a estacionariedade é obtida se (MONTGOMERY, 1997):

$$|\phi_1| < 1$$

Para os processos auto-regressivos de segunda ordem, AR(2), tem estacionariedade assegurada quando (BOX, 1994):

$$\phi_1 + \phi_2 < 1$$

$$\phi_1 - \phi_2 < 1$$

$$|\phi_2| < 1$$

2.4.4 Modelos de Médias Móveis (MA)

Considerando um processo linear $Z_t = \mu + \psi(B)a_t$ e suponha que $\psi_j = 0$, $j > q$; obtemos um processo de médias móveis de ordem q , que denotaremos por $MA(q)$; (MORETTIN & TOLOI, 2006).

O modelo de médias móveis (MA) pode ser representado por uma soma ponderada de ruídos, observados em cada período passado. A representação deste modelo pode ser definida por (MAKRIDAKIS; WHEELWRIGHT; McGEE, 1983);

$$Z_t = \delta + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q},$$

onde:

Z_t é o valor da série no instante t ;

δ representa o termo constante da série;

θ_i corresponde ao i -ésimo parâmetro auto-regressivo, $i = 1, 2, \dots, q$;

a_t é o ruído no instante t ;

Similarmente ao processo auto-regressivo, utiliza-se um operador de média móvel de ordem q para simplificar a representação do modelo, que é dado por:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q. \quad (10)$$

Para a expressão (10) não há restrições sobre os parâmetros θ_j para que o processo seja estacionário. Usando argumento similar ao $AR(p)$, pode se verificar que a condição de invertibilidade para um modelo $MA(q)$ é assegurada quando θ_1 e θ_2 cumprem os mesmos requisitos de estacionariedade de ϕ_1 e ϕ_2 nos processos auto-regressivos (BOX, 1994).

Analicamente, tem-se a inversibilidade para um processo MA(1) quando:

$$|\theta_1| < 1$$

Para um processo MA(2) temos as seguintes condições:

$$\theta_1 + \theta_2 < 1$$

$$\theta_1 - \theta_2 < 1$$

$$|\theta_2| < 1$$

2.4.5 Modelos Auto-Regressivos e de Médias Móveis (ARMA)

A inclusão de ambos os termos auto-regressivos e médias móveis conduz a um modelo mais flexível do que se poderia alcançar com as formas AR puro e MA puro (CHAVES NETO, 2004). Portanto, um modelo que incorpora esses dois termos é o ARMA(p,q), que consegue adequar-se mais facilmente às séries temporais. O modelo é dado por:

$$Z_t = \delta + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} + a_t,$$

ou de forma simplificada:

$$\phi(B)Z_t = \delta + \theta(B)a_t. \quad (11)$$

Analisando a equação (11) é possível verificar que os modelos ARMA relacionam os valores das observações passadas e presente e, também com os erros passados apurados.

A estacionariedade e invertibilidade para o processo auto-regressivo e médias móveis de primeira ordem ARMA(1,1), é assegurada respectivamente, quando $|\phi_1| < 1$ e $|\theta_1| < 1$ (MONTGOMERY, 1997). São as mesmas condições aplicadas individualmente aos modelos AR e MA.

2.4.6 Modelos Auto-Regressivos Integrados de Médias Móveis (ARIMA)

Os modelos ARIMA(p,d,q) englobam todos os modelos vistos anteriormente bem como os processos não estacionários. A metodologia de Box & Jenkins consiste na busca de um modelo ARIMA(p,d,q) que represente o processo estocástico gerador da série temporal, a partir de um modelo ARMA.

Genericamente, um processo ARIMA(p,d,q) pode ser representado pela equação:

$$\omega_t = \phi_1 \omega_{t-1} + \dots + \phi_p \omega_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q},$$

onde

ω_t é o valor da série no instante t, após a diferença ou não;

ϕ_i e θ_j são os parâmetros dos processos auto-regressivos e médias móveis ARMA(p,q), com $i = 1, 2, \dots, p$ e $j = 1, 2, \dots, q$;

a_t , corresponde ao ruído branco que não podem ser explicados pelos modelo;

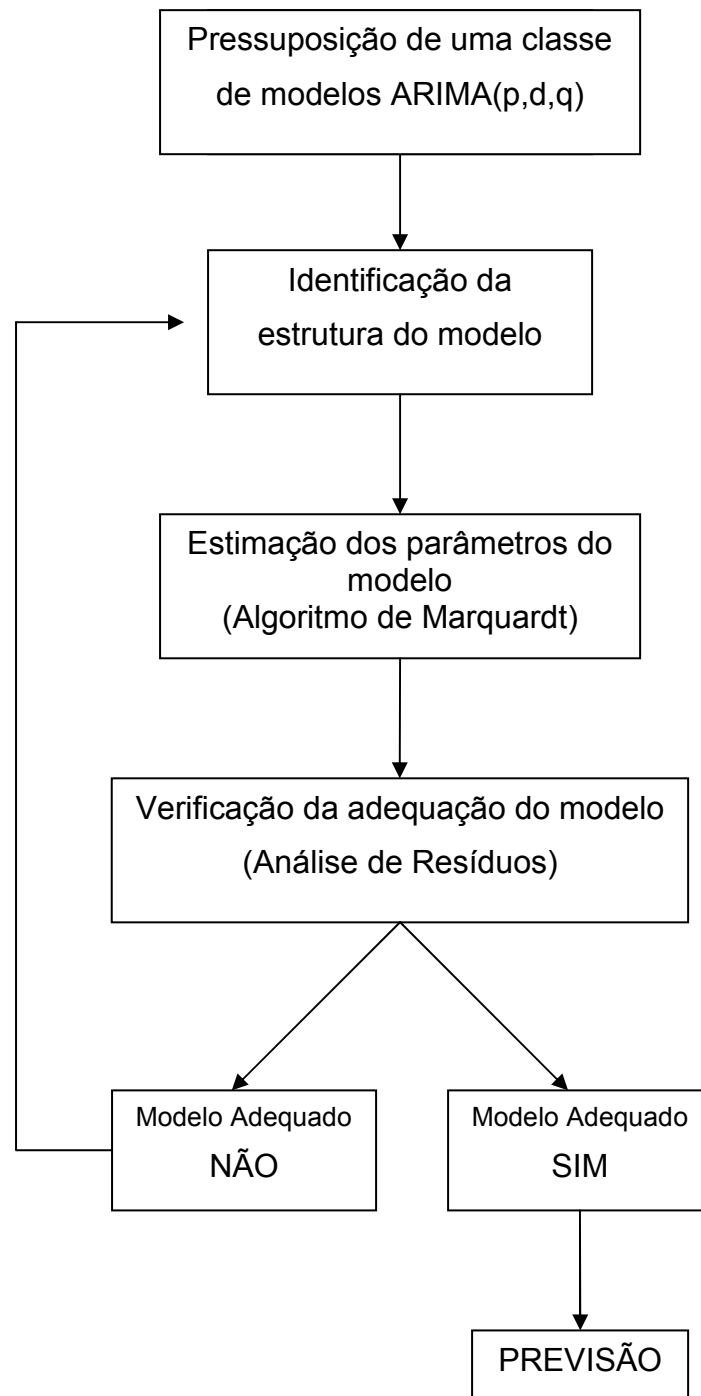
d, equivale ao número de diferenças necessárias para transformar a série em estacionária.

Segundo BOX e JENKINS (1976), a estratégia para a construção de um modelo ARIMA envolve uma abordagem iterativa que pode ser descrita nos seguintes passos:

1. *Identificação* – o objetivo é determinar os valores de p, d e q do modelo ARIMA(p,d,q), esse é uma fase crucial do método.
2. *Estimação* – nesta fase os parâmetros são estimados utilizando o algoritmo de programação não-linear (Marquardt).
3. *Verificação* – pela análise de resíduos, verifica-se se o modelo é bom ou não, caso seja, pode-se passar para a última fase.
4. *Previsão* – depois de identificado o modelo, estimado e verificado, pode-se realizar previsões para períodos futuros.

Através do fluxograma da figura 2 pode-se ter uma visão mais clara das fases dessa metodologia.

FIGURA 2 - FLUXOGRAMA DOS ESTÁGIOS DA METODOLOGIA BOX & JENKINS



FONTE: BOX & JENKINS (1976)

2.4.7 Identificação dos Modelos

A identificação do modelo é a etapa principal dessa metodologia. Segundo CHAVES NETO (1991) trata-se de uma fase crucial do método e aí um erro pode conduzir a resultados desastrosos. A estrutura é identificada por um analista de séries temporais na qual é comparado o correlograma amostral da função de autocorrelação (FAC) e o correlograma da função de autocorrelação parcial (FACP) com os correlogramas teóricos das diversas estruturas ARMA(p,q).

Não há consenso a respeito de qual estratégia seguir na identificação dos modelos (GRANGER & NEWBOLD, 1986). MILLS (1990) define o comportamento esperado dos gráficos da função de autocorrelação e autocorrelação parcial para os modelos ARIMA.

QUADRO 1 - PROPRIEDADES DA FAC E FACP PARA MODELOS ARIMA

MODELO	FAC	FACP
(1,d,0)	Decaimento exponencial ou oscilatório	$b_k = 0$ para $k > 1$
(2,d,0)	Decaimento exponencial ou senoidal	$b_k = 0$ para $k > 2$
(p,d,0)	Decaimento exponencial ou senoidal	$b_k = 0$ para $k > p$
(0,d,1)	$r_k = 0$ para $k > 1$	Prevalece decaimento exponencial
(0,d,2)	$r_k = 0$ para $k > 2$	Prevalece decaimento exponencial ou senoidal
(0,d,q)	$r_k = 0$ para $k > q$	Prevalece combinação linear de decaimento exponencial e/ou senoidal
(1,d,1)	Decaimento exponencial a partir da defasagem 1	Prevalece decaimento exponencial a partir da defasagem 1
(p,d,q)	Decaimento exponencial a partir da defasagem $q - p$	Prevalece decaimento exponencial e/ou senoidal a partir da defasagem $q - p$

FONTE: MILLS (1990)

A análise gráfica da FAC e FACP é uma estratégia muito utilizada pelos analistas de séries temporais para identificação dos modelos, pacotes computacionais, tal como o *Statgraphics* gera esses gráficos automaticamente.

2.5 MÉTODO THETA

2.5.1 Introdução

Os métodos de previsão de séries temporais, classificados como métodos quantitativos, baseiam suas previsões na extrapolação de características de observações passadas, fornecendo previsões muito boas se o futuro apresentar comportamento similar ao passado (WHEELWRIGHT, 1985).

De acordo com WHEELWRIGHT (1985), a maioria dos métodos de previsão de séries temporais se baseia na suposição de que observações passadas contém todas as informações sobre o padrão de comportamento da série temporal e esse padrão é recorrente no tempo. O propósito dos métodos de previsão consiste em distinguir o padrão de qualquer ruído que possa estar contido nas observações e então usar esse padrão para prever os valores futuros da série temporal. SOUZA(1989) classifica os métodos de previsões de séries temporais em: Univariados, Funções de transferência e Multivariados.

Os métodos multivariados abrangem os procedimentos de previsão que associam mais de uma série temporal na efetivação de prognósticos sem, no entanto, qualquer imposição em relação à causalidade entre essas séries.

As metodologias nas quais a série de interesse é explicada não só pelo seu passado histórico, como também por outras séries temporais não-correlatas entre si, são conhecidas como funções de transferência. Esta classe de métodos de previsão envolve, portanto, mais de uma série temporal, com a ressalva de que a relação de causalidade entre estas séries é perfeitamente conhecida.

Os métodos univariados, que compreendem a maior parte dos métodos de previsão de séries temporais, consideram somente uma única série para a realização das predições. As previsões decorrentes da aplicação de métodos univariados podem estar relacionadas apenas com as informações contidas na série histórica de interesse (métodos baseados na estatística clássica) ou também, além de

incorporarem essas informações, consideram outras supostamente relevantes e que não estão contidas na série analisada (métodos baseados na estatística bayesiana).

ASSIMAKOPOULOS & NIKOLOPOULOS (2000) propuseram um novo modelo de previsão univariado chamado de Modelo Theta. Este foi um dos métodos testados na Competição M3 de Makridakis (2000), e apresentou um dos melhores desempenhos. Embora o porquê desta escolha particular de modelo ter sido boa, ainda não foi informado (GOOIJER; HYNDMAN, 2006).

2.5.2 Considerações sobre o Método Theta

O método Theta ou modelo Theta, despertou interesse no meio acadêmico, devido ao seu surpreendente desempenho positivo nas previsões em M3-competition (ASSIMAKOPOULOS & NIKOLOPOULOS, 2000). No entanto, este interesse não foi seguido por um grande número de pesquisadores, com exceção de Hyndman e Billah em 2003.

O modelo theta pode ser entendido segundo a análise de HYNDMAN e BILLAH (2003) como sendo equivalente ao Alisamento Exponencial Simples com “desvio”. Já NIKOLOPOULOS e ASSIMAKOPOULOS (2005) discordam desta abordagem, eles afirmam que o modelo theta é mais genérico que o Alisamento Exponencial Simples porque é uma aproximação da decomposição dos dados que se pode confiar para extrapolação sobre qualquer modelo de previsão.

2.5.3 O Método Theta

O método Theta está baseado na modificação da curvatura local de uma série temporal ajustada sazonalmente através do coeficiente theta (Θ). O coeficiente *theta* é aplicado diretamente na segunda diferença da série temporal (ASSIMAKOPOULOS; NIKOLOPOULOS, 2008). Essa aplicação resulta em uma série chamada *Theta Line* $L(\Theta)$, mantendo a média e a declividade dos dados originais, mas não suas curvaturas.

As observações X_t da série podem ser escritas como:

$$X_i = X_1 + (i-1).(X_2 - X_1) + \left(\sum_{t=2}^{i-1} (i-t).X''_{t+1} \right)$$

onde: $X''_t = X_t - 2X_{t-1} + X_{t-2}$ no tempo t se:

- $\Theta = 0$ a série t é transformada em regressão linear;
- $0 < \Theta < 1$ a série t é deflacionada;
- $\Theta > 1$ a série t é dilatada;

Os pontos Y_i de uma linha *theta* são definidos por:

$$Y_i = Y_1 + (i-1).(Y_2 - Y_1) + \Theta \left(\sum_{t=2}^{i-1} (i-t).X''_{t+1} \right)$$

Desta forma, o problema recai numa minimização de erros quadráticos, ou seja:

$$\min(\sum_i e_i^2) = \min(\sum_i (Y_i - X_i)^2)$$

A formulação geral do método *theta* é:

- Decomposição da série temporal inicial em duas ou mais linhas *theta*;
- Cada uma das linhas *theta* é extrapolada separadamente e as previsões são simplesmente combinadas com pesos iguais.

A formulação mais simples do modelo e também a testada na *M3-Competition* é a decomposição da série temporal em duas linhas *theta*, neste caso as observações da série são decompostas da seguinte forma:

$$X_t = \frac{1}{2}(L(\Theta = 0) + L(\Theta = 2))$$

onde, $L(\Theta = 0)$ é a Regressão Linear dos dados e $L(\Theta = 2)$ é obtida pela seguinte expressão:

$$L(\Theta = 2) = 2.X_t - L(\Theta = 0), \quad (12)$$

onde X_t são os valores da série original.

A $L(\Theta = 0)$ descreve a série como uma tendência linear. A $L(\Theta = 2)$ duplica as curvaturas locais ampliando a ação de curto prazo. Para a extrapolação de $L(\Theta = 2)$ é aplicado o método de Alisamento Exponencial Simples. A previsão final $\hat{X}_t(h)$ para o modelo Theta é obtida combinando as duas linhas theta com pesos iguais, ou seja:

$$\hat{X}_t(h) = \frac{1}{2}(\hat{L}(\Theta = 0)_h + \hat{L}(\Theta = 2)_h), \quad \forall t = 1 \dots n \quad (13)$$

Na prática o modelo pode ser facilmente implementado utilizando a planilha eletrônica do Excel. NIKOLOPOULOS & ASSIMAKOPOULOS (2005) sugerem os seguintes passos para sua implementação:

Passo 0: decomposição sazonal dos dados pelo método clássico multiplicativo caso necessário;

Passo 1: aplicar a regressão linear dos dados $L(\Theta = 0)$, preparar a linha de regressão e as previsões;

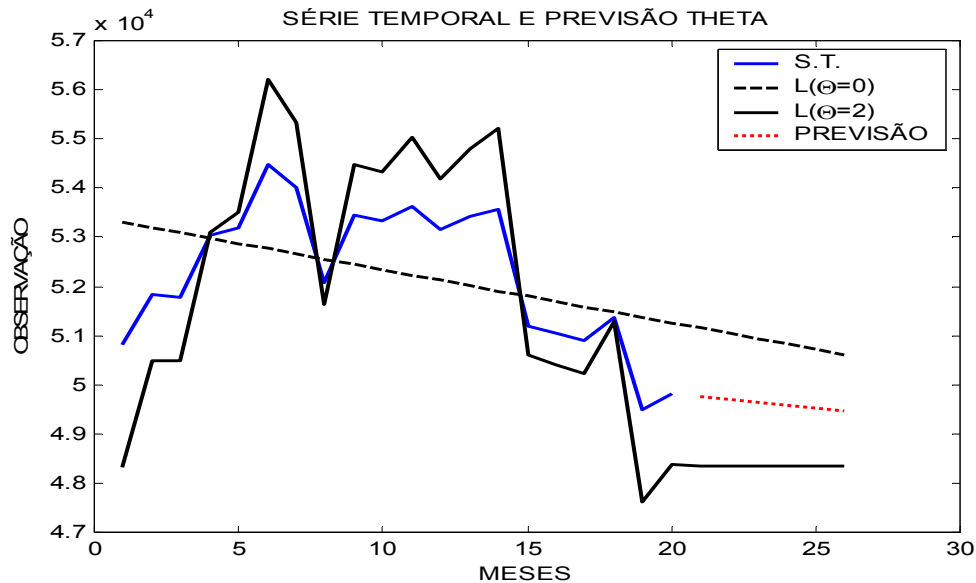
Passo 2: preparar os valores de $L(\Theta = 2)$ utilizando $L(\Theta = 2) = 2X_t - L(\Theta = 0)$;

Passo 3: extrapolar a $L(\Theta = 2)$ com AES (otimizado pelo *solver* do *Microsoft Excel*) ou por outro método mais simples tais como médias móveis;

Passo 4: combinar as previsões com pesos iguais obtidas por regressão linear e AES.

O gráfico 6, a seguir, mostra um exemplo de previsões obtidas pelo Método Theta.

GRÁFICO 6 - SÉRIE DO ÍNDICE BOVESPA E PREVISÃO THETA PELA DECOMPOSIÇÃO EM DUAS $L(\Theta)$



FONTE: O autor (2010)

Para a previsão Theta pode-se extrapolar e combinar várias linhas theta, como por exemplo, em três componentes $L(\Theta = 0)$, $L(\Theta = 0,5)$ e $L(\Theta = 2)$, para a previsão das linhas theta pode ser utilizado qualquer método de previsão (NIKOLOPOULOS & THOMAKOS, 2009). Outra combinação utilizada na competição M3 de Makridakis é de quatro componentes $L(\Theta = 0)$, $L(\Theta = 0,5)$, $L(\Theta = 1,5)$ e $L(\Theta = 2)$, que de longe foi o que obteve melhor resultado (NIKOLOPOULOS & THOMAKOS, 2008). Os valores ótimos de Θ são obtidos por um procedimento de otimização empírico.

2.5.4 Cálculo das Linhas Theta

As linhas theta podem ser calculadas pela seguinte fórmula que foi provada em ASSIMAKOPOULOS & NIKOLOPOULOS (2008):

$$Y_t(\theta) = LRL + \theta \cdot e_{LRL}$$

onde LRL é uma regressão linear dos dados originais da série e e_{LRL} é o resíduo dessa regressão.

Prova:

$$Y_t(\theta) = LRL + \theta.e_{LRL}$$

$$\Leftrightarrow Y_t(\theta) = LRL + \theta.(X_t - LRL)$$

$$\Leftrightarrow Y_t(\theta) = \theta.X_t + LRL - \theta.LRL$$

$$\Leftrightarrow Y_t(\theta) = \theta.X_t + (1 - \theta).LRL$$

Os autores, NIKOLOPOULOS, *et al.* (2009) tentaram encontrar várias formas de otimizar o modelo theta para ver suas limitações. Alguns pesquisadores provaram ter aumentado a precisão nas previsões e outros não conseguiram obter nenhuma melhora.

O método theta em geral é simples e não requer um grande treinamento. De acordo com os resultados da competição M3 de Makridakis e Hibon (2000) o método obteve boas previsões, em séries mensais estacionárias ou com tendência ou sazonalidade. Como desvantagem pode ser citada a falta de intervalos de confiança para as previsões no trabalho de Assimakopoulos e Nikolopoulos (2000).

2.6 TÉCNICA COMPUTACIONALMENTE INTENSIVA “BOOTSTRAP”

2.6.1 Introdução

O “Bootstrap” é um método computacionalmente intensivo desenvolvido por B. Efron para ser usado na estimação da variabilidade de estatísticas. De um modo geral, o “Bootstrap” se refere a uma técnica ou método de simulação que objetiva a estimação por ponto ou por intervalo de confiança de parâmetros de interesse por reamostragem dos dados originais, desde que os métodos clássicos para esse fim sejam assintóticos ou de difícil aplicação.

O método foi criado por Bradley Efron em 1979 e, desde então, tornou-se tão importante que, em estatística, a década de 80 é chamada década do “Bootstrap” (CHAVES NETO, A. 1991). A importância do estudo e aplicação desse

método pode ser percebida pela grande quantidade de artigos, que nas últimas três décadas, estão aparecendo em todo tipo de revistas especializadas.

O “Bootstrap”, como já se disse, é um método computacionalmente intensivo que usa simulação Monte Carlo para estimar erros-padrão e intervalos de confiança. Segundo CHAVES NETO, A. (1991, p36) “O “Bootstrap” é uma técnica estatística não-paramétrica computacionalmente intensiva que permite a avaliação da variabilidade de estatísticas com base nos dados de uma única amostra existente”. Noções como informação prévia, modelos de probabilidade, erro-padrão e intervalos de confiança são destinados a formalizar a incerteza, e assim dar subsídios para a resolução do problema.

Em situações simples as estimativas podem ser obtidas pelo cálculo analítico baseado em um modelo de probabilidade assumido para os dados disponíveis. Mas, em problemas mais complexos, esta abordagem pode se tornar exaustiva e muito difícil, e seus resultados podem ser enganosos. (DAVISON; HINKLEY, 1997). Portanto, o “Bootstrap” é uma metodologia que auxilia em situações onde os métodos padrões falham ou são analiticamente difíceis de examinar.

Entre os métodos computacionalmente intensivos pode-se destacar também o Jackknife, introduzido por Quenouille e Tukey na década de 1950. (CHAVES NETO, A. 1991).

2.6.2 O Método Jackknife

Jackknife é um método computacionalmente intensivo não paramétrico que precedeu o “Bootstrap”. Segundo CHAVES NETO (1991), uma primeira abordagem para esse método foi apresentada em 1949 por Maurice Quenouille (QUENOUILLE, 1949), que consistia em uma técnica para reduzir o vício de um estimador de correlação serial com base na divisão da amostra original em duas semi-amostras. Tal como o “bootstrap” é um método de reamostragem, pois se baseia na construção de subamostras da amostra original.

2.6.3 Obtenção da Amostra Jackknife

Observada uma amostra original de tamanho n , $[x_1, x_2, \dots, x_n]$, o método Jackknife consiste em gerar outras n amostras de tamanho $n-1$, sendo a i -ésima nova amostra ($i = 1, \dots, n$) composta pelas observações que compõem a amostra original, com a exceção da i -ésima observação (MARTINEZ, 2006).

Assim, sendo $[x_1, x_2, \dots, x_n]$ uma amostra de variáveis aleatórias independentes e identicamente distribuídas (i. i. d.) são geradas as amostras $\underline{x}_{(1)}' = (x_2, x_3, \dots, x_n)$, $\underline{x}_{(2)}' = (x_1, x_3, \dots, x_n)$, $\underline{x}_{(3)}' = (x_1, x_2, x_4, \dots, x_n)$, ..., $\underline{x}_{(n-1)}' = (x_1, x_2, \dots, x_{n-2}, x_n)$, $\underline{x}_{(n)}' = (x_1, x_2, \dots, x_{n-1})$. Em cada uma destas amostras geradas, estima-se o parâmetro de interesse θ , por $\hat{\theta}_{(i)} = s(\underline{x}_{(i)})$, $i = 1, 2, \dots, n$, que são as replicações Jackknife da estatística $s(\cdot)$ que é o estimador do parâmetro. Da amostra original \underline{x} , já estava estimado o parâmetro por $\hat{\theta} = s(\underline{x})$, sendo a média das estimativas Jackknife,

$$\hat{\theta}_{(.)} = \frac{\sum_{i=1}^n \hat{\theta}_{(i)}}{n}, \text{ chamada de estimador "Jackknife" do parâmetro } \theta.$$

2.6.4 Método "Bootstrap"

A idéia básica do "Bootstrap", como citado anteriormente é reamostrar um conjunto de observações, diretamente ou via um modelo ajustado, a fim de criar réplicas dos dados, a partir dos quais se pode avaliar a variabilidade de estatísticas, sem o uso de cálculos analíticos.

A reamostragem usada, contudo, não usa a distribuição de probabilidade assumida, ela se baseia em uma distribuição empírica. Aplicando o "bootstrap" obtêm-se novas amostras. Nessa fase usa-se a amostragem com reposição da amostra original. O ponto crucial é a substituição das observações após a amostragem, o que permite a retirada de tantas amostras quantas se deseje. Cada amostra criada pode ser analisada separadamente e os resultados agrupados ao longo da amostra. Um exemplo clássico de uma estimativa "bootstrap", é a de que a melhor estimativa da média é exatamente a média de todas as médias estimadas ao longo da amostra. Este mecanismo também é válido para outras estatísticas além das médias amostrais.

De um modo geral o “Bootstrap” funciona da seguinte maneira:

- Seja $\underline{y}' = [y_1, y_2, \dots, y_n]$ a amostra aleatória disponível da v.a. com função distribuição desconhecida F , $X_i \sim \text{i.i.d. } F$.
- Retira-se dessa amostra uma nova amostra de mesmo tamanho n com reposição. Essa nova amostra obtida com reposição é denominada de “amostra “Bootstrap” e é composta por $\underline{y}' = (y_1^*, y_2^*, \dots, y_n^*)$. Cada observação y_j^* é uma escolha aleatória de (y_1, y_2, \dots, y_n) , fixando-se a probabilidade empírica de $1/n$ para cada observação original ser escolhida. Assim, tem-se a distribuição empírica \hat{F}_n . Tomando-se B amostras “bootstrap” com $i = 1, 2, \dots, B$, tem-se:

$$\begin{array}{ll} y_1^* = [y_{11}^*, y_{12}^*, \dots, y_{1n}^*] & Y_{1i}^* \sim \text{iid } \hat{F}_n \\ y_2^* = [y_{21}^*, y_{22}^*, \dots, y_{2n}^*] & Y_{2i}^* \sim \text{iid } \hat{F}_n \\ \dots\dots\dots & \dots\dots\dots \\ \dots\dots\dots & \dots\dots\dots \\ y_B^* = [y_{B1}^*, y_{B2}^*, \dots, y_{Bn}^*] & Y_{Bi}^* \sim \text{iid } \hat{F}_n \end{array}$$

Estima-se o parâmetro, de interesse, θ por $\hat{\theta}_i^*$ em cada uma das amostras “bootstrap”. A distribuição amostral da estatística $\hat{\theta}_i^*$ $i = 1, 2, \dots, B$ é a distribuição “bootstrap” do estimador $\hat{\theta}_j$ de θ .

É importante salientar que a reamostragem não adiciona informações novas à amostra original. Pode parecer que o “bootstrap” crie dados do nada, entretanto, os dados criados não são utilizados como se fosse reais. O objetivo é verificar como as estatísticas obtidas das reamostras variam em decorrência da amostragem aleatória. A vantagem do “bootstrap” é o resultado da maneira que a amostra é processada. Nos casos de estimação de parâmetros nos quais a distribuição amostral da estatística (estimador) é desconhecida que o “bootstrap” é muito útil.

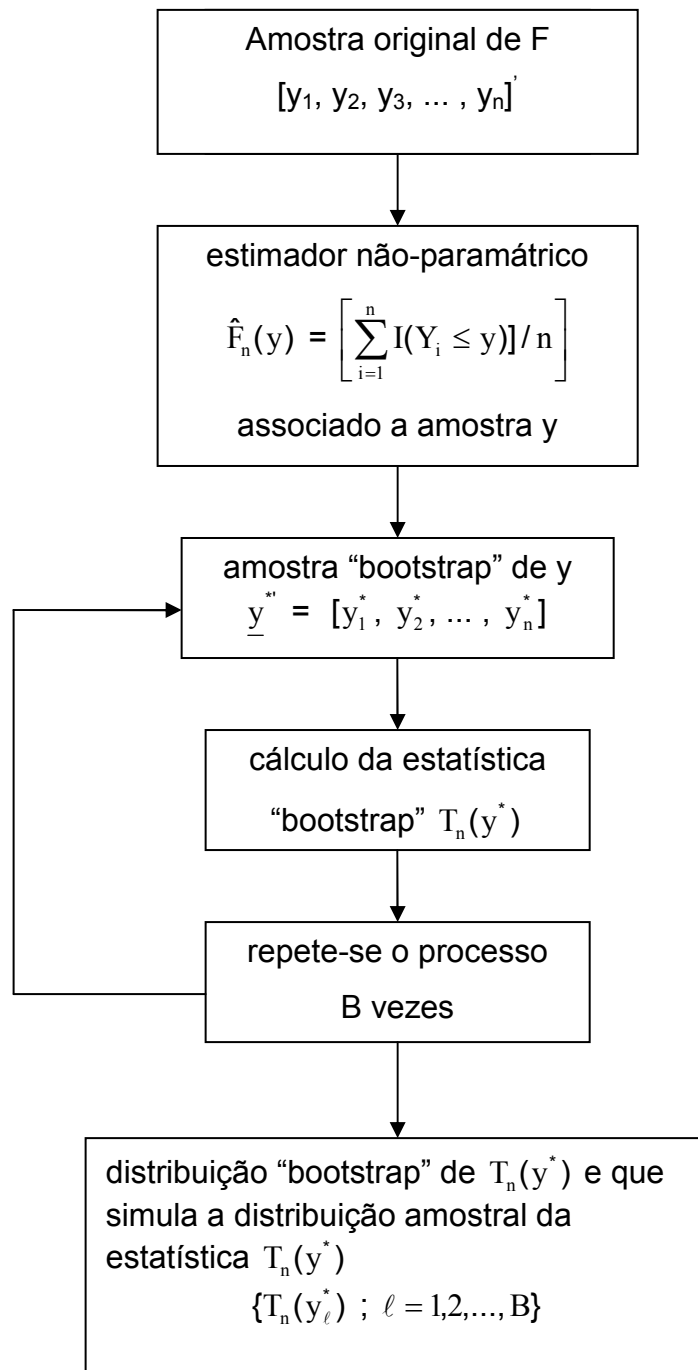
HESTERBERG *et al.* (2003) afirmam que a amostra mestre representa a população da qual ela foi retirada. As reamostras desta amostra representam o que se deve obter quando se retiram muitas amostras da população original. A

distribuição “bootstrap” da estatística, baseada em muitas reamostras, representa uma distribuição amostral dessa estatística.

Para obter dados confiáveis devem ser feitas centenas de reamostras de mesmo tamanho n (mesmo tamanho da amostra original – mestre). A reamostragem é feita com reposição, sempre selecionando os valores de forma aleatória. Depois de geradas essas reamostras, deve-se calcular para cada uma a estatística solicitada no problema.

2.6.5 Obtenção da Amostra “Bootstrap”

O fluxograma da fig. 3 a seguir procura tornar claro o algoritmo de construção da distribuição “bootstrap” descrito anteriormente.

FIGURA 3 - ALGORITMO DA DISTRIBUIÇÃO “BOOTSTRAP” DA ESTATÍSTICA $T_n(x,F)$ 

FONTE: CHAVES NETO (1991)

O procedimento anterior se aplica ao “bootstrap” não paramétrico. Ao se optar pela utilização do “bootstrap” paramétrico, procede-se da mesma forma, com a diferença de que cada amostra “bootstrap” é obtida da distribuição paramétrica que

originou os dados que se tem em mãos, ao invés de reamostrar-se as observações disponíveis.

A aproximação “bootstrap” é válida para a maioria das estatísticas de interesse e que seus estimadores são consistentes. Se $B \rightarrow \infty$ então as estimativas do erro padrão, do enviesamento e do EMQ se igualam as estimativas de máxima verossimilhança (EFRON, 1982). Para o cálculo das estimativas “bootstrap” é preciso um valor de B muito grande, geralmente $B=1000$ produz bons resultados.

2.6.6 Intervalos de Confiança “Bootstrap”

Uma das principais ênfases dadas ao estudo do “bootstrap” é a estimação dos intervalos de confiança (IC). O mais simples dos intervalos “bootstrap” é aquele em se substitui a estimativa usual do erro-padrão, pela estimativa “bootstrap”. Tendo-se, por exemplo, o pivot $\frac{(\hat{\eta}_i - \eta_i)}{s_i}$ com distribuição t_v para o parâmetro η_i de um modelo ARMA(p,q) para séries temporais, o intervalo de nível $1 - \alpha$ para esse parâmetro é:

$$[\hat{\eta}_i - t_v(1 - \alpha/2).s_i \quad ; \quad \hat{\eta}_i + t_v(1 - \alpha/2).s_i]$$

A estatística s_i é o estimador clássico do erro-padrão de $\hat{\eta}_i$. O número de graus de liberdade, v , da distribuição t é dado pela diferença entre o tamanho da série e o número de parâmetros estimados. CHAVES NETO (1991).

O intervalo de confiança “bootstrap” para o parâmetro pode ser obtido simplesmente substituindo-se s_i pela estimativa “bootstrap” s_i^* dada por:

$$s_i^* = \left[\frac{1}{B} \sum_{\ell=1}^B (\hat{\eta}_{i\ell}^* - \hat{\eta}_i)^2 \right]^{\frac{1}{2}},$$

onde:

$\hat{\eta}_{i\ell}^*$: estimativa “bootstrap” do parâmetro η_i na replicação ℓ . Tem-se, portanto:

$$\left[\hat{\eta}_i - t_v(1-\alpha/2).s_i^* \quad ; \quad \hat{\eta}_i + t_v(1-\alpha/2).s_i^* \right]$$

2.6.6.1 Intervalo t-student

Para uma amostra aleatória simples de tamanho n , uma amostra “bootstrap” (EFRON e TIBSHIRANI, 1993) corresponde a uma reamostragem com reposição de mesmo tamanho da amostra original.

Sendo θ o parâmetro de interesse, seu intervalo de confiança “bootstrap” studentizado é obtido calculando as quantidades:

$$K_b^* = \frac{(\hat{\theta}_b^* - \hat{\theta})}{\sqrt{\text{var}(\hat{\theta}_b^*)}},$$

onde:

$\hat{\theta}$: estimativa de θ na amostra original;

$\hat{\theta}_b^*$: é a estimativa de θ na b -ésima reamostra “bootstrap”;

Segundo DAVISON e HINKLEY (1997), tendo em mãos os valores de K_1^* , K_2^* , ..., K_B^* , o α -ésimo quantil da distribuição K^* deve ser estimado por

$$\hat{K}_{B;\alpha} = K_{(B+1)\alpha}^*$$

Assim o intervalo de confiança t-”bootstrap”, é dado por:

$$\left(\hat{\theta} - K_{(B+1)(1-\alpha)}^* \sqrt{\text{var}(\hat{\theta}_b^*)}, \hat{\theta} + K_{(B+1)\alpha}^* \sqrt{\text{var}(\hat{\theta}_b^*)} \right)$$

2.6.6.2 Intervalo “Bootstrap” Residual

Efron (1979) introduziu as aplicações do “bootstrap” em diversas áreas. No caso de modelos lineares, o problema considerado é aquele de estimar a distribuição amostral da estatística $\hat{\beta}$. O modelo geral de regressão é:

$$\underline{\hat{Y}} = X\underline{\hat{\beta}} + \underline{\hat{\varepsilon}},$$

onde:

$\underline{\hat{Y}}$: vetor de respostas com dimensão n;

X: matriz do modelo de ordem nxp;

$\underline{\hat{\beta}}$: vetor de parâmetros com dimensão p;

$\underline{\hat{\varepsilon}}$: vetor de resíduos com dimensão n.

Os erros ε_i , componentes de $\underline{\hat{\varepsilon}}$, são v.a's independentes com distribuição comum F,

$$\varepsilon_i \sim F$$

Em geral F é desconhecida, mas é razoável assumir-se que esta distribuição é centrada em zero e que é Gaussiana, $F = N(0, \sigma_\varepsilon^2)$. (CHAVES NETO, A. 1991)

Os principais passos desse método são:

- 1) Ajustar um modelo de regressão considerando os dados da amostra original e obter o vetor de resíduos $\underline{\hat{\varepsilon}}' = [\hat{\varepsilon}_1, \hat{\varepsilon}_2, \hat{\varepsilon}_3, \dots, \hat{\varepsilon}_n]$.
- 2) Selecionar B amostras aleatórias de tamanho n, dos resíduos, $\underline{\hat{\varepsilon}}$ obtidos no passo (1), utilizando reamostragem com reposição, com probabilidade $1/n$ para cada resíduo selecionado $\underline{\hat{\varepsilon}}_\ell^*$.

$$\underline{\hat{\varepsilon}}_1^* = [\hat{\varepsilon}_{11}^*, \hat{\varepsilon}_{12}^*, \dots, \hat{\varepsilon}_{1n}^*]$$

$$\underline{\hat{\varepsilon}}_2^* = [\hat{\varepsilon}_{21}^*, \hat{\varepsilon}_{22}^*, \dots, \hat{\varepsilon}_{2n}^*]$$

.....

.....

$$\underline{\hat{\varepsilon}}_B^* = [\hat{\varepsilon}_{B1}^*, \hat{\varepsilon}_{B2}^*, \dots, \hat{\varepsilon}_{Bn}^*]$$

$$\hat{\varepsilon}_{\ell j} \sim \text{iid } \hat{F}_n \quad \forall \ell, j$$

- 3) Gerar os novos valores de \bar{Y} pela equação:

$$\underline{Y}_\ell^{*'} = \underline{X}\hat{\underline{\beta}} + \hat{\underline{\varepsilon}}_\ell^*$$

onde $\hat{\underline{\varepsilon}}_\ell^*$ é uma reamostragem dos resíduos obtidos no passo (2);

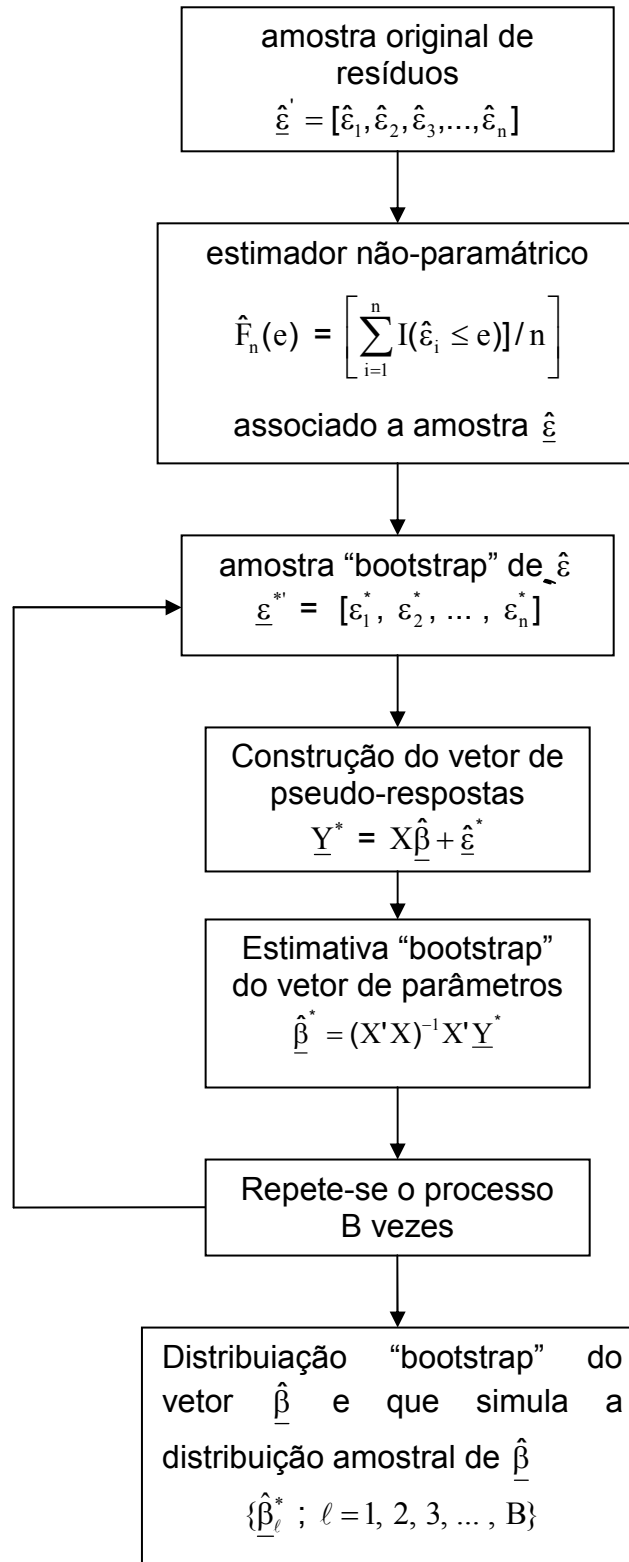
- 4) Ajustar um modelo de regressão pelo método dos mínimos quadrados, utilizando o seguinte modelo:

$$\hat{\underline{\beta}}_\ell^* = [\underline{X}'\underline{X}]^{-1}\underline{X}'\underline{Y}_\ell^*$$

para obter a estimativa de $\underline{\beta}$.

- 5) Repetir os passos (1), (2), , B vezes.

O fluxograma a seguir mostra os passos desse algoritmo.

FIGURA 4 - ALGORITMO DA DISTRIBUIÇÃO “BOOTSTRAP” DA ESTATÍSTICA $\hat{\underline{\beta}}$ 

FONTE: CHAVES NETO (1991)

Depois de B reamostragens, obtém-se a distribuição empírica de $\hat{\beta}^*$ e todas as estatísticas empíricas relacionadas a ela, tais como média “bootstrap” residual e desvio padrão “bootstrap” residual de cada um dos vetores estimados. (MARTINEZ, 2006)

Pode-se obter um intervalo de confiança percentílico para $\hat{\beta}^*$, pela ordenação de $\hat{\beta}_1^* \leq \hat{\beta}_2^* \leq \dots \leq \hat{\beta}_B^*$ e usar $\hat{\beta}_{(B+1)(\alpha^*/2)}^*$ e $\hat{\beta}_{(B+1)(1-\alpha^*)/2}^*$ respectivamente, como limites inferior e superior do intervalo $100(1-\alpha^*)\%$ de confiança para $\hat{\beta}^*$.

3 MATERIAL E MÉTODO

3.1 MATERIAL

Já foi dito no objetivo desta dissertação que, em linhas gerais, o trabalho visa apresentar o Método Theta como uma alternativa para previsões em series temporais e seus respectivos intervalos de confiança, obtidos pelo Método “Bootstrap”. E, também, apresentar o modo de implementar o modelo computacionalmente, fazendo sua validação e comprovando a sua eficiência através de um comparativo com outros métodos automáticos ajustados usando a raiz do erro médio quadrático, RMSE. O Software utilizado para o ajuste de outros métodos foi o *Statgraphics Plus*, versão 2002. Para avaliar a qualidade das previsões obtidas foi utilizado o critério MAPE, ou seja, o erro médio percentual absoluto.

O material que foi trabalhado, ou seja, os dados foram obtidos de séries sintéticas geradas. Para gerar a série sintética usou-se o programa Gerador de Séries Temporais, GST, desenvolvido experimentalmente em linguagem Pascal.

3.1.1 Geração da Série Temporal

Escolheu-se gerar séries com 36 observações e a geração foi feita para modelos AR(1) dado por $Z_t = \delta_t + \phi_1 Z_{t-1} + a_t$, fixando-se o parâmetro auto-regressivo ϕ_1 dentro do espaço paramétrico de estacionariedade, $-1 < \phi_1 < 1$. Considerou-se o valor do termo constante em 45, $\delta_t = 45$, e a variância do ruído foi fixada em $V(a_t) = \sigma_a^2 = 0,2$.

3.2 MÉTODO

Os programas foram desenvolvidos no ambiente Matlab versão 7.1 para atingir os objetivos deste trabalho.

3.2.1 Adaptação dos modelos e implementação computacional

Para aplicação do modelo e obtenção das estimativas dos intervalos de confiança foram adotados os seguintes passos:

1) Escolha dos coeficientes Θ : as previsões para o modelo theta podem ser obtidas de três diferentes combinações de linhas theta:

a) modo clássico com duas linhas theta, $L(\Theta = 0)$ e $L(\Theta = 2)$, extrapoladas por regressão linear (RL) e Alisamento Exponencial Simples (AES) ou Alisamento Exponencial de Brown (AEB) ou Alisamento Exponencial de Holt (AEH), respectivamente;

b) combinação de três linhas theta, $L(\Theta = 0)$, $L(\Theta = ?)$ e $L(\Theta = 2)$, extrapoladas por RL, Alisamento Exponencial de Brown e Alisamento Exponencial Simples e Θ estimado por um procedimento empírico que resulte no menor (EQM);

2) Sazonalidade: não foram feitos testes de sazonalidade.

3) Medidas de desempenho: para avaliar as previsões obtidas foi adotado o critério MAPE, ou seja, o erro médio absoluto percentual.

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right|}{n} \cdot 100$$

4) Comparação com outros métodos de previsão: as previsões obtidas foram comparadas com as obtidas com métodos automáticos ajustados pelo menor RMSE utilizando o Software *Statgraphics*;

3.2.1.1 Método de Previsão Theta

3.2.1.2 Previsões Theta

Dado a série temporal Y_t , ajusta-se à série um modelo de regressão linear, pelo método dos mínimos quadrados ordinários (MQO), obtendo a estimativa de $\hat{\beta}$, e o vetor \hat{Y}_t que será designado como $\underline{L}(\Theta = 0)$.

Para obter as demais linhas theta, substitui-se $\underline{L}(\Theta = 0)$ na equação:

$$\underline{L}(\Theta) = \Theta \underline{Y}_t + (1 - \Theta) \cdot \underline{L}(\Theta = 0) \quad (14)$$

Cada $\underline{L}(\Theta)$ é extrapolada separadamente por um modelo de Alisamento Exponencial (AE) e combinada com pesos iguais. A combinação para o período h, dá a previsão final para o modelo theta.

$$Y_t(h) = \frac{1}{n} (L(\Theta = 0) + L(\Theta)_2 + \dots + L(\Theta)_n) \quad (15)$$

3.2.1.3 Intervalo de Confiança “Bootstrap”

Neste estudo os resíduos são obtidos da amostra \hat{Y}_t , obtida pela combinação com pesos iguais de $\underline{L}(\Theta = 0)^{RL}$ e $\underline{L}(\Theta = 2)^{AE}$ após a aplicação de método de alisamento exponencial. O problema considerado é aquele em que se utiliza de modelos lineares para estimar a distribuição amostral da estatística $\hat{\beta}$ utilizada para estimar $\underline{L}(\Theta = 0)^{RL}$. O modelo geral de regressão adaptado para o contexto de previsão para a regressão linear é:

$$\hat{\underline{L}}(\Theta = 0)^{RL} = X\hat{\beta} + \hat{\varepsilon} ,$$

onde:

$\hat{\underline{L}}(\Theta = 0)^{RL}$: vetor de respostas com dimensão n;

X: matriz do modelo de ordem nxp;

$\hat{\underline{\beta}}$: vetor de parâmetros com dimensão p;

$\hat{\underline{\varepsilon}}$: vetor de resíduos com dimensão n.

Aplicado o modelo regressão linear, o mesmo é utilizado para gerar $\hat{\underline{L}}(\Theta = 2)$ obtido pela equação:

$$\hat{\underline{L}}(\Theta = 2) = 2 \underline{Y}_t - \hat{\underline{L}}(\Theta = 0)^{RL}$$

Para gerar a linha de previsão, ou seja, $\hat{\underline{Y}}_t$, é feita uma combinação com pesos iguais entre $\underline{L}(\Theta = 0)^{RL}$ e $\hat{\underline{L}}(\Theta = 2)^{AE}$, após a aplicação de um método de alisamento exponencial (AE) em $\hat{\underline{L}}(\Theta = 2)$.

$$\hat{\underline{Y}}_t = \frac{1}{2} (\underline{L}(\Theta = 0)^{RL} + \hat{\underline{L}}(\Theta = 2)^{AE})$$

Para efeito de “Bootstrap”, serão utilizados os resíduos obtidos dessa combinação, portanto, tem-se que:

$$\hat{\underline{\varepsilon}} = \underline{Y}_t - \hat{\underline{Y}}_t$$

onde:

\underline{Y}_t : Série original, gerada pelo simulador;

$\hat{\underline{Y}}_t$: Série estimada pelo modelo Theta;

Os passos para a implementação computacional para obter os intervalos de confiança para a previsão h períodos são:

- 1) Ajustar um modelo regressão por mínimos quadrados ordinários (MQO), obtendo a estimativa de $\hat{\beta}$, e o vetor de respostas $\hat{\underline{L}}(\Theta = 0)^{RL}$ e $\hat{\underline{L}}(\Theta = 2)$.
- 2) Aplicar um modelo de Alisamento Exponencial em $\hat{\underline{L}}(\Theta = 2)$, gerando o vetor $\hat{\underline{L}}(\Theta = 2)^{AE}$. Obter o modelo theta pela equação:

$$\hat{\underline{Y}}_t = \frac{1}{2} \left(\hat{\underline{L}}(\Theta = 0)^{RL} + \hat{\underline{L}}(\Theta = 2)^{AE} \right),$$

obtendo-se o vetor de resíduos estimados $\hat{\underline{\varepsilon}}' = [\hat{\varepsilon}_1, \hat{\varepsilon}_2, \hat{\varepsilon}_3, \dots, \hat{\varepsilon}_n]$ que passa a ser considerada a amostra original para efeito do “bootstrap”.

- 3) Selecionar B amostras aleatórias de tamanho n, dos resíduos, $\hat{\underline{\varepsilon}}$ obtidos no passo (1), utilizando reamostragem com reposição, com probabilidade $1/n$ para cada resíduo selecionado $\hat{\underline{\varepsilon}}_\ell^*$.

$$\hat{\underline{\varepsilon}}_1^* = [\hat{\varepsilon}_{11}^*, \hat{\varepsilon}_{12}^*, \dots, \hat{\varepsilon}_{1n}^*]$$

$$\hat{\underline{\varepsilon}}_2^* = [\hat{\varepsilon}_{21}^*, \hat{\varepsilon}_{22}^*, \dots, \hat{\varepsilon}_{2n}^*]$$

.....

.....

$$\hat{\underline{\varepsilon}}_B^* = [\hat{\varepsilon}_{B1}^*, \hat{\varepsilon}_{B2}^*, \dots, \hat{\varepsilon}_{Bn}^*]$$

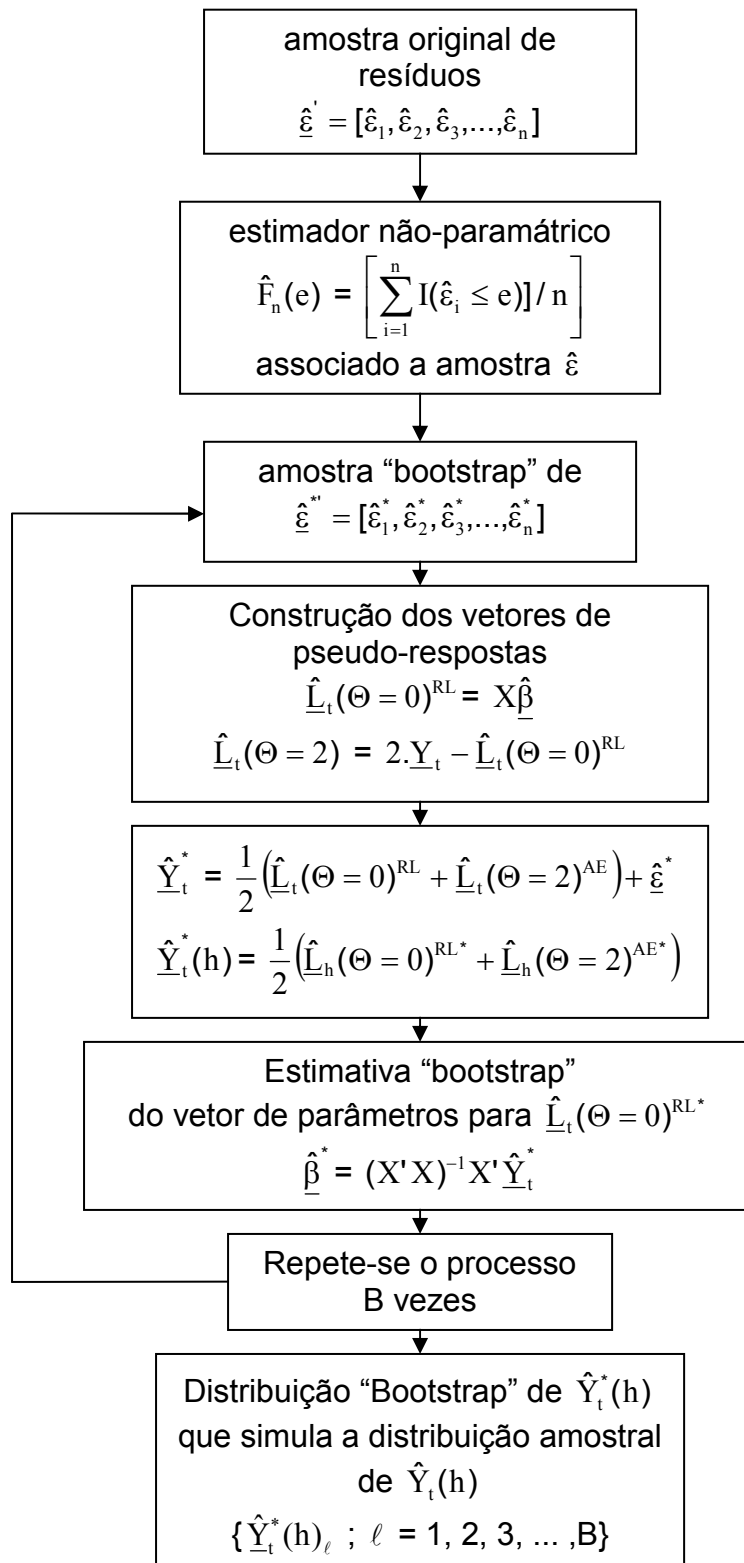
$$\hat{\varepsilon}_{\ell j} \sim \text{iid } \hat{F}_n \quad \forall \ell, j$$

- 4) Gerar as pseudo-séries $\hat{\underline{Y}}_\ell^*$, com cada uma das amostras “bootstrap” pela equação:

$$\hat{\underline{Y}}_\ell^* = \hat{\underline{Y}}_t + \hat{\underline{\varepsilon}}_\ell^*$$

- 5) Ajusta-se novamente o modelo por mínimos quadrados ordinários às pseudo-séries, obtendo a estimativa “bootstrap” de $\hat{\underline{L}}_h(\Theta = 0)^{RL*}$, $\hat{\underline{L}}_h(\Theta = 0)^{AE*}$ e o vetor “bootstrap” para o modelo theta ($\hat{\underline{Y}}^*(h)$).
- 6) Armazenar $\hat{\underline{Y}}^*(h)$ em um vetor Bx1;

O fluxograma a seguir, fig. 5, mostra os passos desse algoritmo.

FIGURA 5 - ALGORITMO DA DISTRIBUIÇÃO “BOOTSTRAP” DE $\hat{Y}(h)$ 

A partir desta distribuição “bootstrap” de $\hat{Y}_t^*(h)$, pode-se calcular o desvio padrão “bootstrap” e os intervalos “bootstrap” para $\hat{Y}_t(h)$. Para se obter um intervalo de confiança percentílico com nível de confiança de 95% para $\hat{Y}_t(h)$, ordena-se em ordem crescente os dados da distribuição “bootstrap”:

$$\hat{Y}_t^*(h)_1 \leq \hat{Y}_t^*(h)_2 \leq \hat{Y}_t^*(h)_3 \leq \dots \leq \hat{Y}_t^*(h)_B$$

e usar $\hat{Y}_t^*(h)_{(B+1)(0.05/2)}$ e $\hat{Y}_t^*(h)_{(B+1)(1-0.05)/2}$ respectivamente, como limites inferior e superior do intervalo 100(1 – 0.05)% de confiança para $\hat{Y}_t(h)$.

4 RESULTADOS E DISCUSSÃO

4.1 PREVISÃO E INTERVALO “BOOTSTRAP” PARA A SÉRIE TEMPORAL

Os resultados da aplicação da metodologia descrita são apresentados da seguinte forma. Primeiro os resultados obtidos pela decomposição da série temporal em duas linhas theta, com os coeficientes theta iguais a 0 e 2, respectivamente e extrapolando os novos valores da série por tendência linear e alisamento exponencial simples, tendência linear e alisamento exponencial linear de Brown e tendência linear e alisamento exponencial de Holt. Indicando os seus RMSE e seus, respectivos, intervalos de confiança.

Em seguida são mostrados os resultados da série temporal decomposta em três linhas theta. Extrapolando as novas linhas theta por tendência linear, alisamento exponencial simples e alisamento exponencial linear de Brown, com os coeficientes theta, respectivamente iguais a 0, 0.5 e 2. E na sequência são obtidos resultados para a série extrapolada por tendência linear, alisamento exponencial linear de Brown e alisamento exponencial simples, com os coeficientes theta, respectivamente iguais a 0, 1.6 e 2, obtendo seus RMSE e intervalos de confiança.

Depois de realizadas as previsões pelo Método Theta, os valores de seus RMSE e MAPE obtidos nas previsões são comparados com os encontrados por outro método de previsão ajustado pelo menor RMSE otimizado pelo Software *Statgraphics*.

4.1.1 Descrição da Série Temporal

A tabela 1, adiante, mostra a série gerada pelo simulador GST, com 36 observações. Os seis últimos valores da série foram guardados para validação e teste de desempenho.

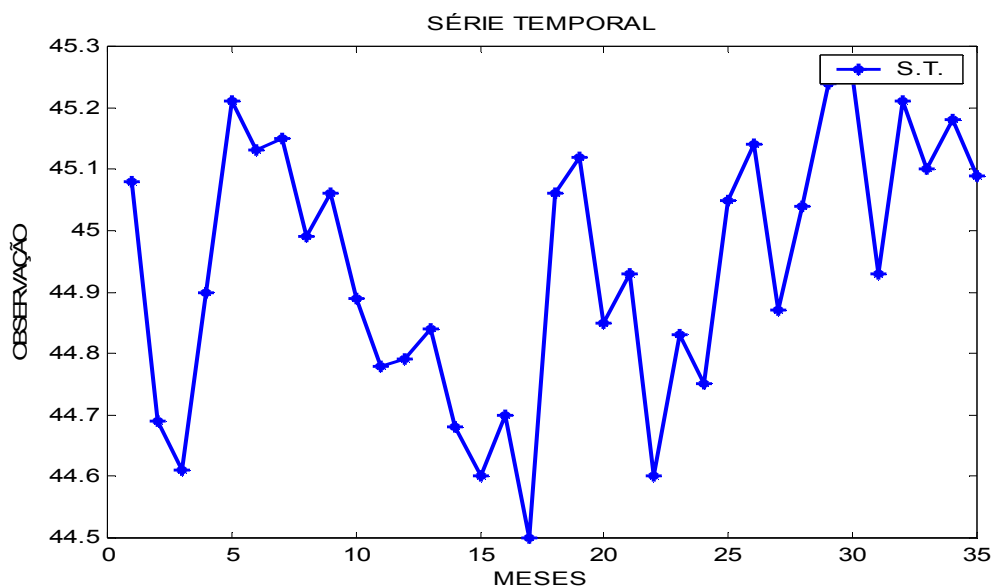
TABELA 1 - SÉRIE TEMPORAL GERADA POR SIMULADOR

Série Temporal			
	Ano1	Ano2	Ano3/teste
Mês 1	45.08	44.84	45.05
Mês 2	44.69	44.68	45.14
Mês 3	44.61	44.60	44.87
Mês 4	44.90	44.70	45.04
Mês 5	45.21	44.50	45.24
Mês 6	45.13	45.06	45.25
Mês 7	45.15	45.12	44.93
Mês 8	44.99	44.85	45.21
Mês 9	45.06	44.93	45.10
Mês 10	44.89	44.60	45.18
Mês 11	44.78	44.83	45.09
Mês 12	44.79	44.75	45.15

FONTE – Dados gerados por simulação utilizando o programa GST.

Abaixo, tem-se a representação gráfica da série Temporal.

GRÁFICO 7 - REPRESENTAÇÃO GRÁFICA DA SÉRIE TEMPORAL



FONTE: O autor (2010)

4.2 RESULTADOS PARA A SÉRIE DECOMPOSTA EM DUAS $L(\Theta)$

O Método Theta é decomposto em duas linhas Theta, ou seja, $L(\Theta=0)$ e $L(\Theta=2)$ e extrapoladas por tendência linear e alisamento exponencial, respectivamente. Os métodos de alisamento exponencial que podem ser utilizados pelo programa são: Alisamento Exponencial Simples (AES), Alisamento Exponencial Linear de Brown (AELB) e Alisamento Exponencial de Holt (AEH). Adiante tem-se um trecho do programa desenvolvido que mostra essa opção:

```

ENTRAR COM A OPÇÃO T=2
*****
*   MÉTODO DE ALISAMENTO EXPONENCIAL   *
*****
*   M = 1, A.E.SIMPLES                  *
*   M = 2, A.E.BROWN                    *
*   M = 3, A.E.HOLT                     *
*****
ENTRAR COM A OPÇÃO M=1
ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS h=6

```

Selecionada a opção M=1 o programa fará a decomposição da série em duas linhas theta e posteriormente a extrapolação de $\hat{L}(\Theta = 2)$ por Alisamento Exponencial Simples. A tabela 2 mostra a série temporal, a decomposição da série com $\hat{L}(\Theta = 0)^{RL}$ descrevendo a série como tendência linear, $\hat{L}(\Theta = 2)$ obtida pela equação (12) e a combinação com pesos iguais obtida pela expressão (13).

TABELA 2 - DECOMPOSIÇÃO DA SÉRIE EM DUAS LINHAS THETA

Período	Série T	$\hat{\underline{L}}(\Theta = 0)^{RL}$	$\hat{\underline{L}}(\Theta = 2)$	Combinação
1	45.08	44.8603	45.2997	45.0800
2	44.69	44.8638	44.5162	44.6900
3	44.61	44.8673	44.3527	44.6100
4	44.90	44.8708	44.9292	44.9000
5	45.21	44.8743	45.5457	45.2100
6	45.13	44.8778	45.3822	45.1300
7	45.15	44.8813	45.4187	45.1500
8	44.99	44.8848	45.0952	44.9900
9	45.06	44.8883	45.2317	45.0600
10	44.89	44.8918	44.8882	44.8900
11	44.78	44.8953	44.6647	44.7800
12	44.79	44.8988	44.6812	44.7900
13	44.84	44.9023	44.7777	44.8400
14	44.68	44.9058	44.4542	44.6800
15	44.60	44.9093	44.2907	44.6000
16	44.70	44.9127	44.4873	44.7000
17	44.50	44.9162	44.0838	44.5000
18	45.06	44.9197	45.2003	45.0600
19	45.12	44.9232	45.3168	45.1200
20	44.85	44.9267	44.7733	44.8500
21	44.93	44.9302	44.9298	44.9300
22	44.60	44.9337	44.2663	44.6000
23	44.83	44.9372	44.7228	44.8300
24	44.75	44.9407	44.5593	44.7500
25	45.05	44.9442	45.1558	45.0500
26	45.14	44.9477	45.3323	45.1400
27	44.87	44.9512	44.7888	44.8700
28	45.04	44.9547	45.1253	45.0400
29	45.24	44.9582	45.5218	45.2400
30	45.25	44.9617	45.5383	45.2500

FONTE: O autor (2010)

4.2.1 PT21 - Tendência Linear e Alisamento Exponencial Simples (AES)

Na tabela 3 encontra-se os valores de $\hat{\underline{L}}(\Theta = 0)^{RL}$ extrapolados por Tendência Linear, $\hat{\underline{L}}(\Theta = 2)^{AES}$ extrapolados por Alisamento Exponencial Simples, a combinação estimada $\hat{\underline{Y}}_t$ para a previsão Theta e a amostra de resíduos $\hat{\underline{\varepsilon}}$.

TABELA 3 - EXTRAPOLAÇÃO DAS $L(\Theta)$ POR TENDÊNCIA LINEAR E ALISAMENTO EXPONENCIAL SIMPLES E OBTENÇÃO DA AMOSTRA DE RESÍDUOS

Período t	Série T Y_t	$\hat{L}(\Theta = 0)^{RL}$	$\hat{L}(\Theta = 2)^{AES}$	Combinação \hat{Y}_t	Resíduo $\hat{\varepsilon}$
1	45.08	44.8603	45.2997	45.0800	0.0000
2	44.69	44.8638	45.2997	45.0817	-0.3917
3	44.61	44.8673	44.6886	44.7779	-0.1679
4	44.90	44.8708	44.4266	44.6487	0.2513
5	45.21	44.8743	44.8186	44.8465	0.3635
6	45.13	44.8778	45.3857	45.1318	-0.0018
7	45.15	44.8813	45.3830	45.1321	0.0179
8	44.99	44.8848	45.4108	45.1478	-0.1578
9	45.06	44.8883	45.1647	45.0265	0.0335
10	44.89	44.8918	45.2170	45.0544	-0.1644
11	44.78	44.8953	44.9605	44.9279	-0.1479
12	44.79	44.8988	44.7298	44.8143	-0.0243
13	44.84	44.9023	44.6919	44.7971	0.0429
14	44.68	44.9058	44.7589	44.8323	-0.1523
15	44.60	44.9093	44.5213	44.7153	-0.1153
16	44.70	44.9127	44.3415	44.6271	0.0729
17	44.50	44.9162	44.4552	44.6857	-0.1857
18	45.06	44.9197	44.1655	44.5426	0.5174
19	45.12	44.9232	44.9726	44.9479	0.1721
20	44.85	44.9267	45.2411	45.0839	-0.2339
21	44.93	44.9302	44.8762	44.9032	0.0268
22	44.60	44.9337	44.9180	44.9259	-0.3259
23	44.83	44.9372	44.4097	44.6734	0.1566
24	44.75	44.9407	44.6539	44.7973	-0.0473
25	45.05	44.9442	44.5801	44.7622	0.2878
26	45.14	44.9477	45.0291	44.9884	0.1516
27	44.87	44.9512	45.2656	45.1084	-0.2384
28	45.04	44.9547	44.8937	44.9242	0.1158
29	45.24	44.9582	45.0744	45.0163	0.2237
30	45.25	44.9617	45.4234	45.1925	0.0575

FONTE: O autor (2010)

4.2.1.1 Previsões e Intervalo “Bootstrap” para PT21

A tabela 4, adiante, mostra os seis valores da série guardados para teste de desempenho, as previsões para $L_h(\Theta = 0)^{RL}$ e $L_h(\Theta = 2)^{AES}$, a previsão para o Método Theta $\hat{Y}_t(h)$, a avaliação da qualidade da previsão de acordo com o critério MAPE, EQM e RMSE.

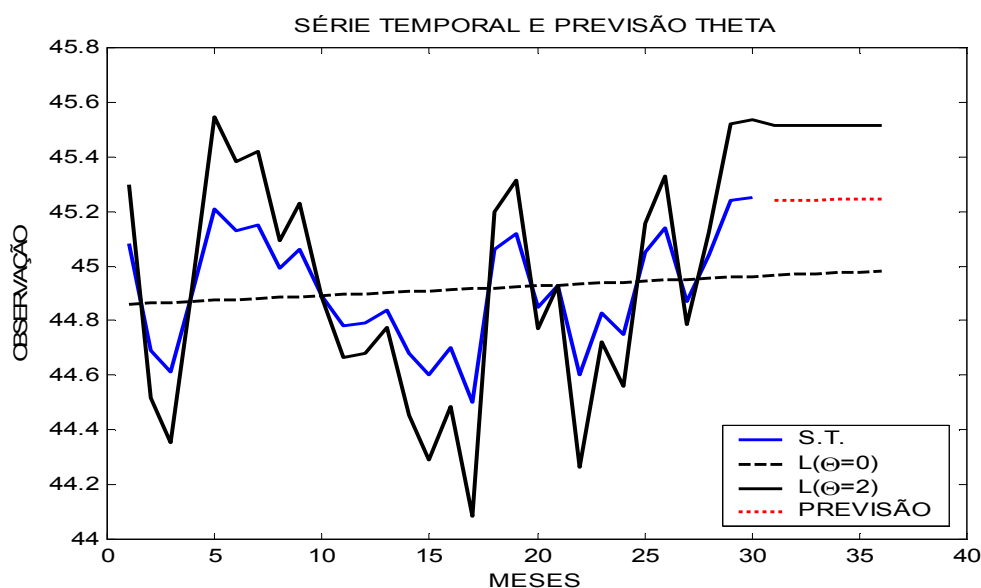
TABELA 4 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM)

Período (h)	Observado (teste)	$\underline{L}_h(\Theta = 0)^{RL}$	$\underline{L}_h(\Theta = 2)^{AES}$	$\hat{Y}_t(h)$	MAPE	EQM
1	44.93	44.9652	45.5130	45.2391	0.6880	0.0955
2	45.21	44.9687	45.5130	45.2409	0.0682	0.0010
3	45.10	44.9722	45.5130	45.2426	0.3162	0.0203
4	45.18	44.9757	45.5130	45.2443	0.1424	0.0041
5	45.09	44.9792	45.5130	45.2461	0.3462	0.0244
6	45.15	44.9826	45.5130	45.2478	0.2167	0.0096
Média					0.2963%	0.0258
RMSE						0.1606

FONTE: O autor (2010)

O RMSE para a série decomposta em duas linhas e pela extrapolação por RL e AES, respectivamente, é igual a 0,1606. A medida de desempenho médio de acordo com o critério MAPE é de 0,2963%. O gráfico 8, adiante, mostra a série temporal analisada, a linha de regressão linear de alisamento exponencial e as previsões para o método theta.

GRÁFICO 8 - SÉRIE TEMPORAL, $L(\Theta=0)^{RL}$, $L(\Theta=2)^{AES}$ E PREVISÃO THETA



FONTE: O autor (2010)

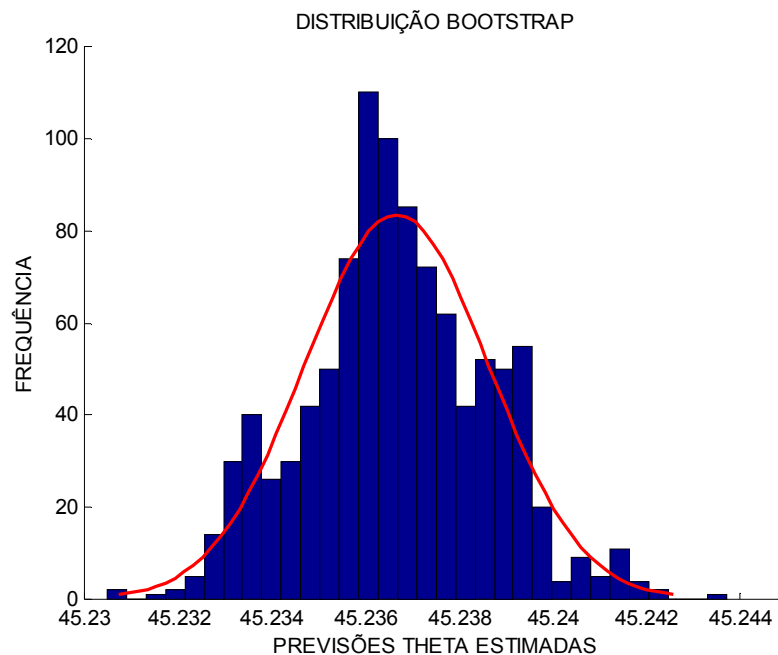
A tabela 5, adiante, mostra os valores observados, as previsões para o Método Theta, Desvio Padrão “bootstrap”, EQM “bootstrap” e os Limites Inferiores e Superiores de Confiança.

TABELA 5 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO “BOOTSTRAP” E O INTERVALO DE CONFIANÇA (NÍVEL DE 5% E 1000 REPLICAÇÕES)

Período (h)	Observado $\underline{Y}_t(h)$	P.Theta $\hat{\underline{Y}}_t(h)$	Desvio- Padrão “bootstrap”	EQM “bootstrap”	Limite Inf. 95%	Limite Sup. 95%
1	44.93	45.2391	0.0020	0.09×10^{-4}	45.2329	45.2406
2	45.21	45.2409	0.0040	0.37×10^{-4}	45.2286	45.2439
3	45.10	45.2426	0.0062	0.86×10^{-4}	45.2228	45.2488
4	45.18	45.2443	0.0085	1.49×10^{-4}	45.2200	45.2522
5	45.09	45.2461	0.0108	2.42×10^{-4}	45.2141	45.2561
6	45.15	45.2478	0.0125	3.12×10^{-4} 1.39×10^{-4}	45.2123	45.2600

FONTE: O autor (2010)

Analisando a tabela 5, nota-se que o EQM “bootstrap” relativo as previsões obtidas, aparecem em ordem crescente, significando que a cada horizonte de previsão os limites de confiança aparecem em um intervalo mais amplo. Tornando as previsões menos confiáveis a medida que o horizonte de previsão aumenta. O intervalo “Bootstrap” para a previsão no horizonte de h períodos à frente para a série temporal, aplicou-se o Método “Bootstrap” com um número de replicações $B = 1000$. O histograma das estimativas “bootstrap” da previsão $\hat{Y}_t^*(h)$, para a primeira previsão, fig. 6 a seguir, mostra a distribuição “bootstrap” para $\hat{Y}_t^*(1)$. No gráfico percebe-se um nível alto de simetria, o que sugere um modelo Gaussiano para esses valores obtidos.

FIGURA 6 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT21 PARA $\hat{Y}_T^*(1)$ 

FONTE: O autor (2010)

4.2.2 PT22 - Tendência Linear e Alisamento Exponencial Linear de Brown (AELB)

```

ENTRAR COM A OPÇÃO T=2
*****
* MÉTODO DE ALISAMENTO EXPONENCIAL *
*****
* M = 1, A.E.SIMPLES *
* M = 2, A.E.BROWN *
* M = 3, A.E.HOLT *
*****
ENTRAR COM A OPÇÃO M=2
ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS h=6

```

Selecionada a opção M=2 o programa fará a decomposição da série em duas linhas theta e posteriormente a extrapolação de $\hat{\underline{L}}(\Theta = 2)$ por Alisamento Exponencial Linear de Brown. Na tabela 6 encontra-se os valores de $\hat{\underline{L}}(\Theta = 0)^{RL}$ extrapolados por Tendência Linear, $\hat{\underline{L}}(\Theta = 2)^{AELB}$ extrapolados por Alisamento Exponencial Linear de Brown, a combinação estimada $\hat{\underline{Y}}_t$ para a previsão Theta e a amostra de resíduos $\hat{\underline{\varepsilon}}$.

TABELA 6 - EXTRAPOLAÇÃO DAS $L(\Theta)$ POR TENDÊNCIA LINEAR E ALISAMENTO EXPONENCIAL LINEAR DE BROWN E OBTENÇÃO DA AMOSTRA DE RESÍDUOS

Período t	Série T Y_t	$\hat{L}(\Theta = 0)^{RL}$	$\hat{L}(\Theta = 2)^{AELB}$	Combinação \hat{Y}_t	Resíduo $\hat{\epsilon}$
1	45.08	44.8603	45.2997	45.0800	0.0000
2	44.69	44.8638	44.8230	44.8434	-0.1534
3	44.61	44.8673	44.5138	44.6906	-0.0806
4	44.90	44.8708	44.7516	44.8112	0.0888
5	45.21	44.8743	45.2462	45.0603	0.1497
6	45.13	44.8778	45.3529	45.1153	0.0147
7	45.15	44.8813	45.3981	45.1397	0.0103
8	44.99	44.8848	45.2160	45.0504	-0.0604
9	45.06	44.8883	45.2168	45.0525	0.0075
10	44.89	44.8918	45.0169	44.9543	-0.0643
11	44.78	44.8953	44.7930	44.8441	-0.0641
12	44.79	44.8988	44.7142	44.8065	-0.0165
13	44.84	44.9023	44.7490	44.8256	0.0144
14	44.68	44.9058	44.5714	44.7386	-0.0586
15	44.60	44.9093	44.3920	44.6506	-0.0506
16	44.70	44.9127	44.4413	44.6770	0.0230
17	44.50	44.9162	44.2261	44.5712	-0.0712
18	45.06	44.9197	44.8084	44.8641	0.1959
19	45.12	44.9232	45.1459	45.0345	0.0855
20	44.85	44.9267	44.9355	44.9311	-0.0811
21	44.93	44.9302	44.9218	44.9260	0.0040
22	44.60	44.9337	44.5223	44.7280	-0.1280
23	44.83	44.9372	44.6250	44.7811	0.0489
24	44.75	44.9407	44.5900	44.7653	-0.0153
25	45.05	44.9442	44.9325	44.9384	0.1116
26	45.14	44.9477	45.1923	45.0700	0.0700
27	44.87	44.9512	44.9594	44.9553	-0.0853
28	45.04	44.9547	45.0491	45.0019	0.0381
29	45.24	44.9582	45.3410	45.1496	0.0904
30	45.25	44.9617	45.4752	45.2184	0.0316

FONTE: O autor (2010)

4.2.2.1 Previsões e Intervalo “Bootstrap” para PT22

A tabela 7, adiante, mostra os seis valores da série guardados para teste de desempenho, as previsões para $L_h(\Theta = 0)^{RL}$ e $L_h(\Theta = 2)^{AELB}$, a previsão para o Método Theta $\hat{Y}_t(h)$, a avaliação da qualidade da previsão de acordo com o critério MAPE, EQM e RMSE.

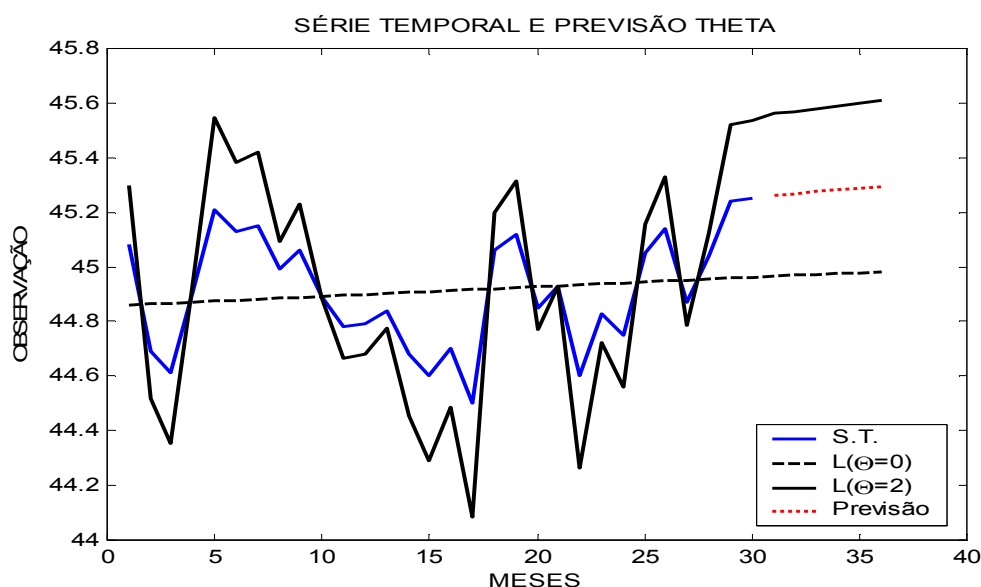
TABELA 7 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM)

Período (h)	Observado (teste)	$\underline{L}_h(\Theta = 0)^{RL}$	$\underline{L}_h(\Theta = 2)^{AELB}$	$\hat{Y}_t(h)$	MAPE	EQM
1	44.93	44.9652	45.5603	45.2628	0.7406	0.1107
2	45.21	44.9687	45.5698	45.2692	0.1310	0.0035
3	45.10	44.9722	45.5793	45.2757	0.3896	0.0309
4	45.18	44.9757	45.5887	45.2822	0.2262	0.0104
5	45.09	44.9792	45.5982	45.2887	0.4406	0.0395
6	45.15	44.9826	45.6076	45.2951	0.3215	0.0211
Média					0.3749%	0.0360
RMSE						0.1897

FONTE: O autor (2010)

O RMSE para a série decomposta em duas linhas e pela extrapolação por RL e AELB, respectivamente é igual a 0.1897. A medida de desempenho médio de acordo com o critério MAPE é de 0.3749%. O gráfico 9, adiante, mostra a série temporal analisada, a linha de regressão linear, alisamento exponencial e as previsões para o método theta.

GRÁFICO 9 - SÉRIE TEMPORAL, $L(\Theta=0)^{RL}$, $L(\Theta=2)^{AELB}$ E PREVISÃO THETA



FONTE: O autor (2010)

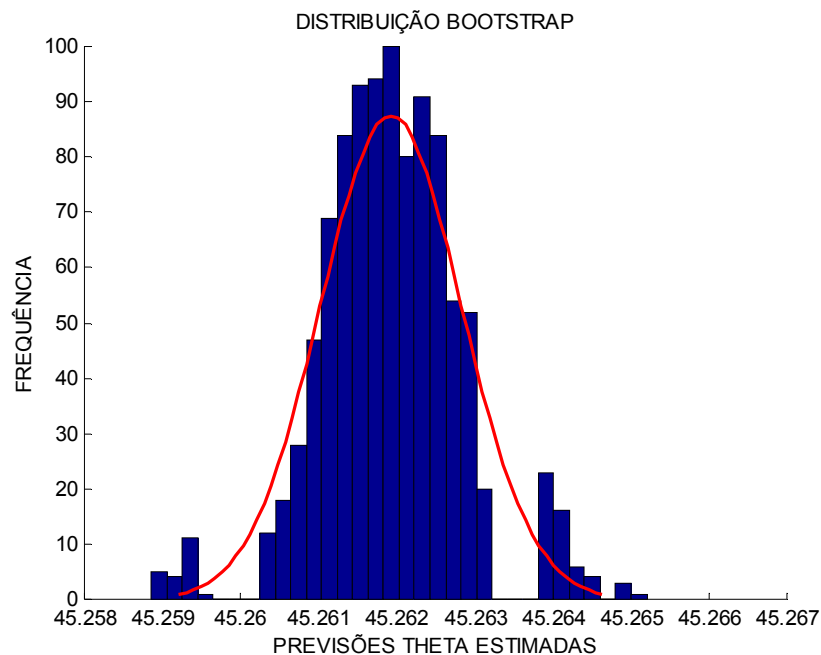
A tabela 8, adiante, mostra os valores observados, as previsões para o Método Theta, Desvio Padrão “bootstrap”, EQM “bootstrap” e os Limites Inferiores e Superiores de Confiança.

TABELA 8 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO “BOOTSTRAP” E O INTERVALO DE CONFIANÇA (NÍVEL DE 5% E 1000 REPLICAÇÕES)

Período (h)	Observado $\underline{Y}_t(h)$	P.Theta $\hat{\underline{Y}}_t(h)$	Desvio- Padrão “bootstrap”	EQM “Bootstrap”	Limite Inf. 95%	Limite Sup. 95%
1	44.93	45.2628	0.0009	0.16×10^{-5}	45.2603	45.2642
2	45.21	45.2692	0.0018	0.59×10^{-5}	45.2644	45.2715
3	45.10	45.2757	0.0026	1.39×10^{-5}	45.2664	45.2792
4	45.18	45.2822	0.0035	2.31×10^{-5}	45.2723	45.2871
5	45.09	45.2887	0.0042	3.38×10^{-5}	45.2770	45.2943
6	45.15	45.2951	0.0050	5.06×10^{-5} 2.15×10^{-5}	45.2810	45.3013

FONTE: O autor (2010)

Analisando a tabela 8, nota-se que o EQM “bootstrap” relativo as previsões obtidas, aparecem em ordem crescente, significando que a cada horizonte de previsão os limites de confiança aparecem em um intervalo mais amplo. Tornando as previsões menos confiáveis a medida que o horizonte de previsão aumenta. O intervalo “Bootstrap” para a previsão no horizonte de h períodos à frente para a série temporal, aplicou-se o Método “Bootstrap” com um número de replicações $B = 1000$. O histograma das estimativas “bootstrap” da previsão $\hat{Y}_t^*(h)$, para a primeira previsão, fig. 7 a seguir, mostra a distribuição “bootstrap” para $\hat{Y}_t^*(1)$. No gráfico percebe-se um nível alto de simetria, o que sugere um modelo Gaussiano para esses valores obtidos.

FIGURA 7 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT22 PARA $\hat{Y}_T^*(1)$ 

FONTE: O autor (2010)

4.2.3 PT23 - Tendência Linear e Alisamento Exponencial de Holt (AEH)

```

ENTRAR COM A OPÇÃO T=2
*****
*   MÉTODO DE ALISAMENTO EXPONENCIAL   *
*****
*   M = 1, A.E.SIMPLES                  *
*   M = 2, A.E.BROWN                    *
*   M = 3, A.E.HOLT                      *
*****
ENTRAR COM A OPÇÃO M=3
ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS h=6

```

Selecionada a opção M=3 o programa fará a decomposição da série em duas linhas theta e posteriormente a extrapolação de $\hat{L}(\Theta = 2)$ por Alisamento Exponencial de Holt. Na tabela 9 encontra-se os valores de $\hat{L}(\Theta = 0)^{RL}$ extrapolados por Tendência Linear, $\hat{L}(\Theta = 2)^{AEH}$ extrapolados por Alisamento Exponencial de Holt, a combinação estimada \hat{Y}_t para a previsão Theta e a amostra de resíduos $\hat{\varepsilon}$.

TABELA 9 - EXTRAPOLAÇÃO DAS $L(\Theta)$ POR TENDÊNCIA LINEAR E ALISAMENTO EXPONENCIAL DE HOLT E OBTENÇÃO DA AMOSTRA DE RESÍDUOS

Período t	Série T Y_t	$\hat{L}(\Theta = 0)^{RL}$	$\hat{L}(\Theta = 2)^{AEH}$	Combinação \hat{Y}_t	Resíduo $\hat{\varepsilon}$
1	45.08	44.8603	45.2997	45.0800	0.0000
2	44.69	44.8638	45.1962	45.0300	-0.3400
3	44.61	44.8673	44.4570	44.6622	-0.0522
4	44.90	44.8708	44.1779	44.5244	0.3756
5	45.21	44.8743	44.6867	44.7805	0.4295
6	45.13	44.8778	45.3824	45.1301	-0.0001
7	45.15	44.8813	45.3735	45.1274	0.0226
8	44.99	44.8848	45.4071	45.1459	-0.1559
9	45.06	44.8883	45.1120	45.0002	0.0598
10	44.89	44.8918	45.1846	45.0382	-0.1482
11	44.78	44.8953	44.8820	44.8886	-0.1086
12	44.79	44.8988	44.6193	44.7590	0.0310
13	44.84	44.9023	44.5927	44.7475	0.0925
14	44.68	44.9058	44.6882	44.7970	-0.1170
15	44.60	44.9093	44.4133	44.6613	-0.0613
16	44.70	44.9127	44.2157	44.5642	0.1358
17	44.50	44.9162	44.3724	44.6443	-0.1443
18	45.06	44.9197	44.0366	44.4782	0.5818
19	45.12	44.9232	45.0253	44.9743	0.1457
20	44.85	44.9267	45.3323	45.1295	-0.2795
21	44.93	44.9302	44.8777	44.9040	0.0260
22	44.60	44.9337	44.9302	44.9319	-0.3319
23	44.83	44.9372	44.3193	44.6283	0.2017
24	44.75	44.9407	44.6300	44.7854	-0.0354
25	45.05	44.9442	44.5438	44.7440	0.3060
26	45.14	44.9477	45.0876	45.0177	0.1223
27	44.87	44.9512	45.3584	45.1548	-0.2848
28	45.04	44.9547	44.8960	44.9254	0.1146
29	45.24	44.9582	45.1151	45.0366	0.2034
30	45.25	44.9617	45.5263	45.2440	0.0060

FONTE: O autor (2010)

4.2.3.1 Previsões e Intervalo “Bootstrap” para PT23

A tabela 10, adiante, mostra os seis valores da série guardados para teste de desempenho, a previsão para $L_h(\Theta = 0)^{RL}$ e $L_h(\Theta = 2)^{AEH}$, a previsão para o Método Theta $\hat{Y}_t(h)$, a avaliação da qualidade da previsão de acordo com o critério MAPE, EQM e RMSE.

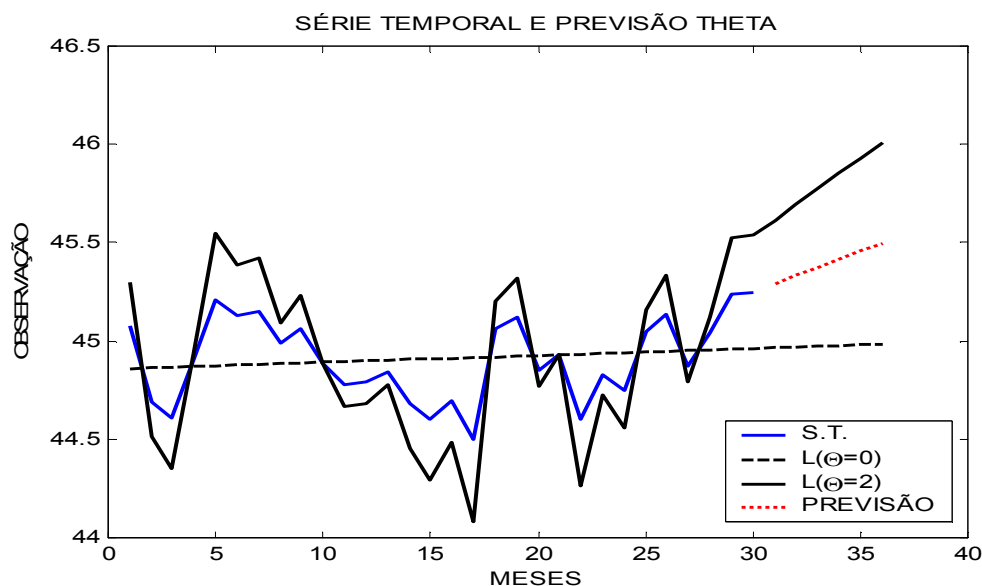
TABELA 10 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM)

Período (h)	Observado (teste)	$\underline{L}_h(\Theta = 0)^{RL}$	$\underline{L}_h(\Theta = 2)^{AEH}$	$\hat{Y}_t(h)$	MAPE	EQM
1	44.93	44.9652	45.6153	45.2902	0.8017	0.1298
2	45.21	44.9687	45.6944	45.3315	0.2688	0.0148
3	45.10	44.9722	45.7735	45.3728	0.6049	0.0744
4	45.18	44.9757	45.8526	45.4141	0.5182	0.0548
5	45.09	44.9792	45.9317	45.4554	0.8104	0.1335
6	45.15	44.9826	46.0108	45.4967	0.7679	0.1202
Média					0.6287%	0.0879
RMSE						0.2964

FONTE: O autor (2010)

O RMSE para a série decomposta em duas linhas e pela extrapolação por RL e AEH, respectivamente é igual a 0.2964. A medida de desempenho médio de acordo com o critério MAPE é de 0.6287%. O gráfico 10, adiante, mostra a série temporal analisada, a linha de regressão linear, alisamento exponencial e as previsões para o método theta.

GRÁFICO 10 - SÉRIE TEMPORAL, $L(\Theta=0)^{RL}$, $L(\Theta=2)^{AEH}$ E PREVISÃO THETA



FONTE: O autor (2010)

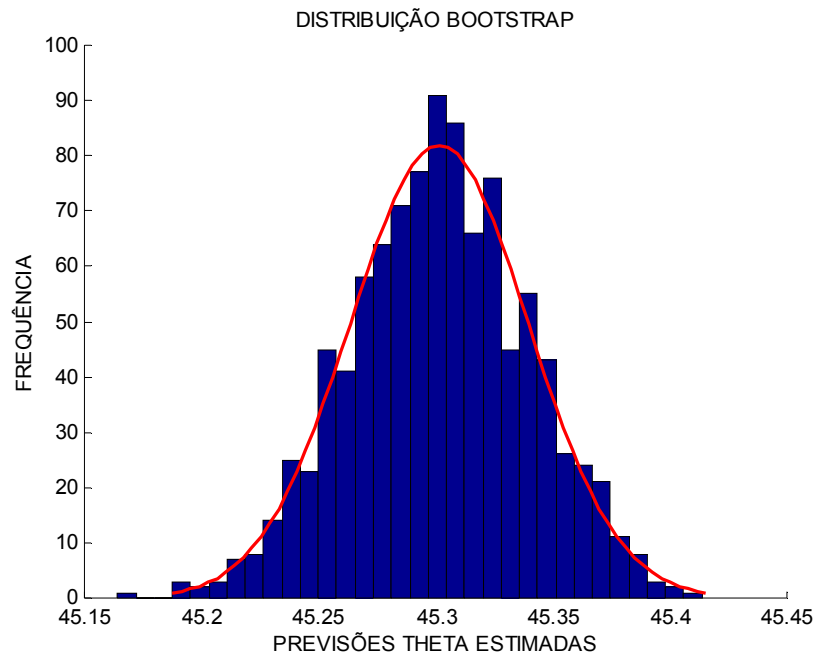
A tabela 11, adiante, mostra os valores observados, as previsões para o Método Theta, Desvio Padrão “bootstrap”, EQM “bootstrap” e os Limites Inferiores e Superiores de Confiança.

TABELA 11 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO “BOOTSTRAP” E O INTERVALO DE CONFIANÇA (NÍVEL DE 5% E 1000 REPLICAÇÕES)

Período (h)	Observado $\underline{Y}_t(h)$	P.Theta $\hat{\underline{Y}}_t(h)$	Desvio- Padrão “bootstrap”	EQM “Bootstrap”	Limite Inf. 95%	Limite Sup. 95%
1	44.93	45.2902	0.0391	1.72×10^{-3}	45.2227	45.3784
2	45.21	45.3315	0.0417	1.90×10^{-3}	45.2587	45.4272
3	45.10	45.3728	0.0448	2.10×10^{-3}	45.3022	45.4724
4	45.18	45.4141	0.0468	2.31×10^{-3}	45.3243	45.5193
5	45.09	45.4554	0.0505	2.49×10^{-3}	45.3622	45.5639
6	45.15	45.4967	0.0506	2.83×10^{-3}	45.4139	45.6097
Média				2.23×10^{-3}		

FONTE: O autor (2010)

Analisando a tabela 11, nota-se que o EQM “bootstrap” relativo as previsões obtidas, aparecem em ordem crescente, significando que a cada horizonte de previsão os limites de confiança aparecem em um intervalo mais amplo. Tornando as previsões menos confiáveis a medida que o horizonte de previsão aumenta. O intervalo “Bootstrap” para a previsão no horizonte de h períodos à frente para a série temporal, aplicou-se o Método “Bootstrap” com um número de replicações $B = 1000$. O histograma das estimativas “bootstrap” da previsão $\hat{Y}_t^*(h)$, para a primeira previsão, fig. 8 a seguir, mostra a distribuição “bootstrap” para $\hat{Y}_t^*(1)$. No gráfico percebe-se um nível alto de simetria, o que sugere um modelo Gaussiano para esses valores obtidos.

FIGURA 8 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT23 PARA $\hat{Y}_t^*(1)$ 

FONTE: O autor (2010)

4.3 RESULTADOS PARA A SÉRIE DECOMPOSTA EM TRÊS $L(\Theta)$

4.3.1 PT31 - Tendência Linear, Alisamento Exponencial Simples (AES) e Alisamento Exponencial de Brown (AEB)

O Método Theta é decomposto em três linhas Theta, ou seja, $L(\Theta=0)$, $L(\Theta=0,5)$ e $L(\Theta=2)$ e extrapoladas por tendência linear (RL), Alisamento exponencial Simples (AES) e Alisamento exponencial Linear de Brown (AELB), respectivamente. O terceiro valor de Θ e método de Alisamento exponencial utilizado foi otimizado pelo menor EQM, resultando em $\Theta=0,5$. O programa utilizado para a otimização de Θ encontra-se no Apêndice E. Adiante tem-se um trecho do programa desenvolvido que mostra essas opções:

```

ENTRAR COM A OPÇÃO T=3
*****
*      quantidades de linhas theta      *
*****
*      T = 2, L(0) e L(2)                *
*      T = 3, L(0), L(?) e L(2)          *
*****
ENTRAR COM A OPÇÃO T=3
ENTRAR COM A ORDEM DE AE, SE 1 = AELB e AES, SE 2 = AES E AELB ORD=2
ENTRAR COM O VALOR DO COEFICIENTE THETA  $\theta=0.5$ 

```


ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS $h=6$

Selecionada a opção $T=3$ e $\Theta=0.5$ o programa fará a decomposição da série em três linhas theta e posteriormente a extrapolação de $\hat{L}(\Theta = 0.5)$ por Alisamento Exponencial Linear de Brown e $\hat{L}(\Theta = 2)$ por Alisamento Exponencial Simples. A tabela 12 mostra a série temporal, a decomposição da série com $\hat{L}(\Theta = 0)^{RL}$ descrevendo a série como tendência linear, $\hat{L}(\Theta = 0.5)$ obtida pela equação (14), $\hat{L}(\Theta = 2)$ obtida pela equação (12) e a combinação com pesos iguais obtida pela expressão (15).

TABELA 12 - DECOMPOSIÇÃO DA S.T. EM TRÊS LINHAS THETA

Período	Série T	$\hat{L}(\Theta = 0)^{RL}$	$\hat{L}(\Theta = 0.5)$	$\hat{L}(\Theta = 2)$	Combinação
1	45.08	44.8603	44.9702	45.2997	45.0434
2	44.69	44.8638	44.7769	44.5162	44.7190
3	44.61	44.8673	44.7387	44.3527	44.6529
4	44.90	44.8708	44.8854	44.9292	44.8951
5	45.21	44.8743	45.0422	45.5457	45.1541
6	45.13	44.8778	45.0039	45.3822	45.0880
7	45.15	44.8813	45.0156	45.4187	45.1052
8	44.99	44.8848	44.9374	45.0952	44.9725
9	45.06	44.8883	44.9741	45.2317	45.0314
10	44.89	44.8918	44.8909	44.8882	44.8903
11	44.78	44.8953	44.8376	44.6647	44.7992
12	44.79	44.8988	44.8444	44.6812	44.8081
13	44.84	44.9023	44.8711	44.7777	44.8504
14	44.68	44.9058	44.7929	44.4542	44.7176
15	44.60	44.9093	44.7546	44.2907	44.6515
16	44.70	44.9127	44.8064	44.4873	44.7355
17	44.50	44.9162	44.7081	44.0838	44.5694
18	45.06	44.9197	44.9899	45.2003	45.0366
19	45.12	44.9232	45.0216	45.3168	45.0872
20	44.85	44.9267	44.8884	44.7733	44.8628
21	44.93	44.9302	44.9301	44.9298	44.9300
22	44.60	44.9337	44.7669	44.2663	44.6556
23	44.83	44.9372	44.8836	44.7228	44.8479
24	44.75	44.9407	44.8454	44.5593	44.7818
25	45.05	44.9442	44.9971	45.1558	45.0324
26	45.14	44.9477	45.0438	45.3323	45.1079
27	44.87	44.9512	44.9106	44.7888	44.8835
28	45.04	44.9547	44.9973	45.1253	45.0258
29	45.24	44.9582	45.0991	45.5218	45.1930
30	45.25	44.9617	45.1058	45.5383	45.2019

FONTE: O autor (2010)

Na tabela 13 encontra-se os valores de $\hat{L}(\Theta = 0)^{RL}$ extrapolado por Tendência Linear, $\hat{L}(\Theta = 0.5)^{AES}$ extrapolado por Alisamento Exponencial Simples, $\hat{L}(\Theta = 2)^{AELB}$ extrapolado por Alisamento Exponencial Linear de Brown, a combinação estimada \hat{Y}_t para a previsão Theta e a amostra de resíduos $\hat{\varepsilon}$.

TABELA 13 - EXTRAPOLAÇÃO DAS $L(\Theta)$ POR TENDÊNCIA LINEAR, ALISAMENTO EXPONENCIAL SIMPLES E ALISAMENTO EXPONENCIAL LINEAR DE BROWN E OBTENÇÃO DA AMOSTRA DE RESÍDUOS

Período t	Série T Y_t	$\hat{L}(\Theta = 0)^{RL}$	$\hat{L}(\Theta = 0.5)^{AES}$	$\hat{L}(\Theta = 2)^{AELB}$	Combinação \hat{Y}_t	Resíduo $\hat{\varepsilon}$
1	45.08	44.8603	44.9702	45.2997	45.0434	0.0366
2	44.69	44.8638	44.9702	44.8230	44.8857	-0.1957
3	44.61	44.8673	44.8175	44.5138	44.7329	-0.1229
4	44.90	44.8708	44.7552	44.7516	44.7925	0.1075
5	45.21	44.8743	44.8581	45.2462	44.9929	0.2171
6	45.13	44.8778	45.0035	45.3529	45.0781	0.0519
7	45.15	44.8813	45.0038	45.3981	45.0944	0.0556
8	44.99	44.8848	45.0132	45.2160	45.0380	-0.0480
9	45.06	44.8883	44.9533	45.2168	45.0194	0.0406
10	44.89	44.8918	44.9698	45.0169	44.9595	-0.0695
11	44.78	44.8953	44.9075	44.7930	44.8652	-0.0852
12	44.79	44.8988	44.8523	44.7142	44.8217	-0.0317
13	44.84	44.9023	44.8460	44.7490	44.8324	0.0076
14	44.68	44.9058	44.8659	44.5714	44.7810	-0.1010
15	44.60	44.9093	44.8082	44.3920	44.7032	-0.1032
16	44.70	44.9127	44.7659	44.4413	44.7066	-0.0066
17	44.50	44.9162	44.7979	44.2261	44.6468	-0.1468
18	45.06	44.9197	44.7270	44.8084	44.8184	0.2416
19	45.12	44.9232	44.9347	45.1459	45.0013	0.1187
20	44.85	44.9267	45.0034	44.9355	44.9552	-0.1052
21	44.93	44.9302	44.9125	44.9218	44.9215	0.0085
22	44.60	44.9337	44.9264	44.5223	44.7942	-0.1942
23	44.83	44.9372	44.8004	44.6250	44.7875	0.0425
24	44.75	44.9407	44.8661	44.5900	44.7989	-0.0489
25	45.05	44.9442	44.8497	44.9325	44.9088	0.1412
26	45.14	44.9477	44.9662	45.1923	45.0354	0.1046
27	44.87	44.9512	45.0275	44.9594	44.9794	-0.1094
28	45.04	44.9547	44.9352	45.0491	44.9796	0.0604
29	45.24	44.9582	44.9843	45.3410	45.0945	0.1455
30	45.25	44.9617	45.0750	45.4752	45.1706	0.0794

FONTE: O autor (2010)

4.3.1.1 Previsões e Intervalo “Bootstrap” para PT31

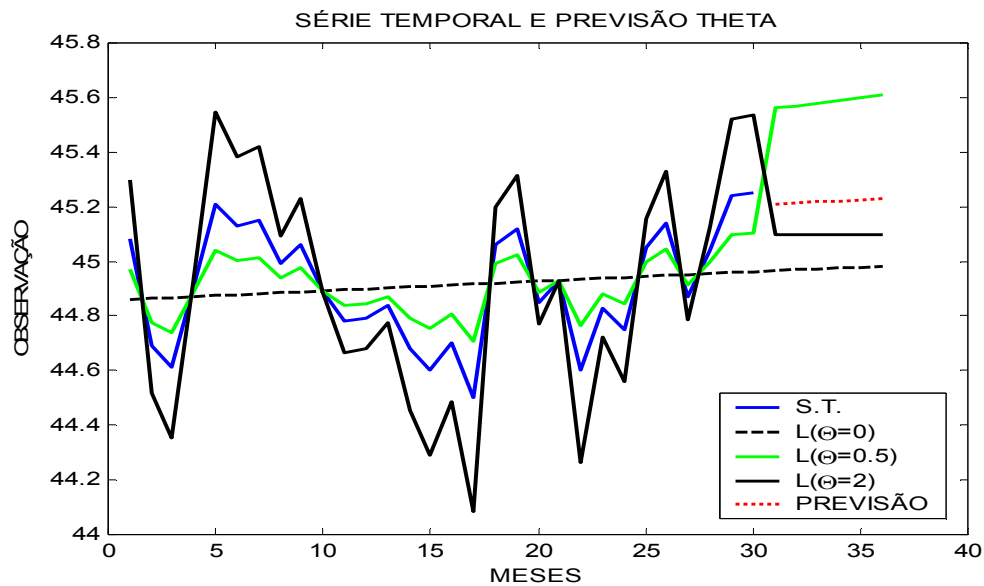
A tabela 14, adiante, mostra os seis valores da série guardados para teste de desempenho, as previsões para $L_h(\Theta = 0)^{RL}$, $L_h(\Theta = 0.5)^{AES}$ e $L_h(\Theta = 2)^{AELB}$, a previsão para o Método Theta $\hat{Y}_t(h)$, a avaliação da qualidade da previsão de acordo com o critério MAPE, EQM e RMSE.

TABELA 14 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM)

Per. (h)	Obs. (teste)	$L_h(\Theta = 0)^{RL}$	$L_h(\Theta=0.5)^{AE}$	$L_h(\Theta=2)^{AEL}$	$\hat{Y}_t(h)$	MAPE	EQM
1	44.93	44.9652	45.0994	45.5603	45.208	0.6194	0.077
2	45.21	44.9687	45.0994	45.5698	45.212	0.0058	0.000
3	45.10	44.9722	45.0994	45.5793	45.216	0.2593	0.013
4	45.18	44.9757	45.0994	45.5887	45.221	0.0913	0.001
5	45.09	44.9792	45.0994	45.5982	45.225	0.3006	0.018
6	45.15	44.9826	45.0994	45.6076	45.229	0.1769	0.006
Méd. RMSE						0.242%	0.019 0.1378

FONTE: O autor (2010)

O RMSE para a série decomposta em três linhas e pela extrapolação por RL, AES e AELB, respectivamente é igual a 0.1378. A medida de desempenho médio de acordo com o critério MAPE é de 0.242%. O gráfico 11, adiante, mostra a série temporal analisada, a linha de regressão linear, de alisamento exponencial e as previsões para o método theta.

GRÁFICO 11 - SÉRIE TEMPORAL, $L(\Theta=0)^{RL}$, $L(\Theta=0.5)^{AES}$, $L(\Theta=2)^{AELB}$ E PREVISÃO THETA

FONTE: O autor (2010)

A tabela 15, adiante, mostra os valores observados, as previsões para o Método Theta, Desvio Padrão “bootstrap”, EQM “bootstrap” e os Limites Inferiores e Superiores de Confiança.

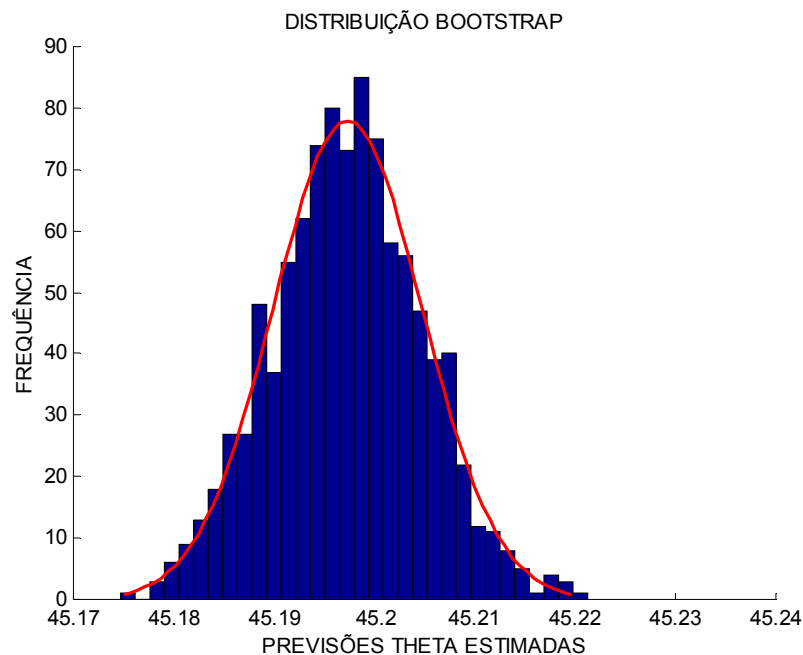
TABELA 15 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO “BOOTSTRAP” E O INTERVALO DE CONFIANÇA (NÍVEL DE 5% E 1000 REPLICAÇÕES)

Período (h)	Observado $\underline{Y}_t(h)$	P.Theta $\hat{\underline{Y}}_t(h)$	Desvio- Padrão “bootstrap”	EQM “Bootstrap”	Limite Inf. 95%	Limite Sup. 95%
1	44.9300	45.2083	0.0074	1.80×10^{-4}	45.1814	45.2110
2	45.2100	45.2126	0.0083	2.17×10^{-4}	45.1844	45.2180
3	45.1000	45.2169	0.0093	2.68×10^{-4}	45.1861	45.2220
4	45.1800	45.2212	0.0103	3.26×10^{-4}	45.1868	45.2271
5	45.0900	45.2256	0.0108	3.79×10^{-4}	45.1883	45.2306
6	45.1500	45.2299	0.0119	4.31×10^{-4}	45.1893	45.2362
				3.00×10^{-4}		

FONTE: O autor (2010)

Analisando a tabela 15, nota-se que o EQM “bootstrap” relativo as previsões obtidas, aparecem em ordem crescente, significando que a cada horizonte de previsão os limites de confiança aparecem em um intervalo mais amplo. Tornando as previsões menos confiáveis a medida que o horizonte de previsão aumenta. O intervalo “Bootstrap” para a previsão no horizonte de h períodos à frente para a série temporal, aplicou-se o Método “Bootstrap” com um número de replicações $B = 1000$. O histograma das estimativas “bootstrap” da previsão $\hat{Y}_t^*(h)$, para a primeira previsão, fig. 9 a seguir, mostra a distribuição “bootstrap” para $\hat{Y}_t^*(1)$. No gráfico percebe-se um nível alto de simetria, o que sugere um modelo Gaussiano para esses valores obtidos.

FIGURA 9 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT31 PARA $\hat{Y}_T^*(1)$



FONTE: O autor (2010)

4.3.2 PT32 - Tendência Linear, Alisamento Exponencial de Brown (AEB) e Alisamento Exponencial Simples (AES)

O Método Theta é decomposto em três linhas Theta, ou seja, $L(\Theta=0)$, $L(\Theta=1,6)$ e $L(\Theta=2)$ e extrapoladas por tendência linear (RL), Alisamento exponencial Linear de Brown (AELB) e Alisamento exponencial Simples (AES), respectivamente. O terceiro valor de Θ e método de Alisamento exponencial utilizado foi otimizado pelo menor EQM, resultando em $\Theta=1,6$. O programa utilizado para a otimização de Θ encontra-se no Apêndice E. Adiante tem-se um trecho do programa desenvolvido que mostra essas opções:

```

                                ENTRAR COM A OPÇÃO T=3
*****
*      quantidades de linhas theta      *
*****
*      T = 2, L(0) e L(2)                *
*      T = 3, L(0), L(?) e L(2)         *
*****

                                ENTRAR COM A OPÇÃO T=3
ENTRAR COM A ORDEM DE AE, SE 1 = AELB e AES, SE 2 = AES E AELB ORD=1
                                ENTRAR COM O VALOR DO COEFICIENTE THETA  $\theta=1.6$ 
                                ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS h=6

```

Selecionada a opção $T=3$ e $\Theta=1,6$ o programa fará a decomposição da série em três linhas theta e posteriormente a extrapolação de $\hat{L}(\Theta=1,6)$ por Alisamento Exponencial Linear de Brown e $\hat{L}(\Theta=2)$ por Alisamento Exponencial Simples. A tabela 16 mostra a série temporal, a decomposição da série com $\hat{L}(\Theta=0)^{RL}$ descrevendo a série como tendência linear, $\hat{L}(\Theta=1,6)$ obtida pela equação (14), $\hat{L}(\Theta=2)$ obtida pela equação (12) e a combinação com pesos iguais obtida pela expressão (15).

TABELA 16 - DECOMPOSIÇÃO DA SÉRIE TEMPORAL EM TRÊS LINHAS THETA E COMBINAÇÃO

Período	Série T	$\hat{\underline{L}}(\Theta = 0)^{RL}$	$\hat{\underline{L}}(\Theta = 1.6)$	$\hat{\underline{L}}(\Theta = 2)$	Combinação
1	45.08	44.8603	45.2118	45.2997	45.1239
2	44.69	44.8638	44.5857	44.5162	44.6552
3	44.61	44.8673	44.4556	44.3527	44.5585
4	44.90	44.8708	44.9175	44.9292	44.9058
5	45.21	44.8743	45.4114	45.5457	45.2771
6	45.13	44.8778	45.2813	45.3822	45.1804
7	45.15	44.8813	45.3112	45.4187	45.2037
8	44.99	44.8848	45.0531	45.0952	45.0110
9	45.06	44.8883	45.1630	45.2317	45.0943
10	44.89	44.8918	44.8889	44.8882	44.8896
11	44.78	44.8953	44.7108	44.6647	44.7569
12	44.79	44.8988	44.7247	44.6812	44.7682
13	44.84	44.9023	44.8026	44.7777	44.8275
14	44.68	44.9058	44.5445	44.4542	44.6348
15	44.60	44.9093	44.4144	44.2907	44.5381
16	44.70	44.9127	44.5724	44.4873	44.6575
17	44.50	44.9162	44.2503	44.0838	44.4168
18	45.06	44.9197	45.1442	45.2003	45.0881
19	45.12	44.9232	45.2381	45.3168	45.1594
20	44.85	44.9267	44.8040	44.7733	44.8347
21	44.93	44.9302	44.9299	44.9298	44.9300
22	44.60	44.9337	44.3998	44.2663	44.5333
23	44.83	44.9372	44.7657	44.7228	44.8086
24	44.75	44.9407	44.6356	44.5593	44.7119
25	45.05	44.9442	45.1135	45.1558	45.0712
26	45.14	44.9477	45.2554	45.3323	45.1785
27	44.87	44.9512	44.8213	44.7888	44.8538
28	45.04	44.9547	45.0912	45.1253	45.0571
29	45.24	44.9582	45.4091	45.5218	45.2964
30	45.25	44.9617	45.4230	45.5383	45.3077

FONTE: O autor (2010)

Na tabela 17 encontra-se os valores de $\hat{\underline{L}}(\Theta = 0)^{RL}$ extrapolado por Tendência Linear, $\hat{\underline{L}}(\Theta = 1.6)^{AELB}$ extrapolado por Alisamento Exponencial Linear de Brown, $\hat{\underline{L}}(\Theta = 2)^{AES}$ extrapolado por Alisamento Exponencial Simples, a combinação estimada $\hat{\underline{Y}}_t$ para a previsão Theta e a amostra de resíduos $\hat{\underline{\epsilon}}_t$.

TABELA 17 - EXTRAPOLAÇÃO DAS $L(\Theta)$ POR TENDÊNCIA LINEAR, ALISAMENTO EXPONENCIAL LINEAR DE BROWN E ALISAMENTO EXPONENCIAL SIMPLES E OBTENÇÃO DA AMOSTRA DE RESÍDUOS

Período t	Série T \underline{Y}_t	$\hat{L}(\Theta = 0)^{RL}$	$\hat{L}(\Theta = 1.6)^{AELB}$	$\hat{L}(\Theta = 2)^{AES}$	Combinação \hat{Y}_t	Resíduo $\hat{\varepsilon}$
1	45.08	44.8603	45.2118	45.2997	45.1239	-0.0439
2	44.69	44.8638	44.8309	45.2997	44.9981	-0.3081
3	44.61	44.8673	44.5841	44.6886	44.7133	-0.1033
4	44.90	44.8708	44.7750	44.4266	44.6908	0.2092
5	45.21	44.8743	45.1714	44.8186	44.9548	0.2552
6	45.13	44.8778	45.2575	45.3857	45.1737	-0.0437
7	45.15	44.8813	45.2943	45.3830	45.1862	-0.0362
8	44.99	44.8848	45.1494	45.4108	45.1483	-0.1583
9	45.06	44.8883	45.1507	45.1647	45.0679	-0.0079
10	44.89	44.8918	44.9915	45.2170	45.0334	-0.1434
11	44.78	44.8953	44.8130	44.9605	44.8896	-0.1096
12	44.79	44.8988	44.7507	44.7298	44.7931	-0.0031
13	44.84	44.9023	44.7793	44.6919	44.7912	0.0488
14	44.68	44.9058	44.6378	44.7589	44.7675	-0.0875
15	44.60	44.9093	44.4951	44.5213	44.6419	-0.0419
16	44.70	44.9127	44.5352	44.3415	44.5965	0.1035
17	44.50	44.9162	44.3638	44.4552	44.5784	-0.0784
18	45.06	44.9197	44.8303	44.1655	44.6385	0.4215
19	45.12	44.9232	45.1009	44.9726	44.9989	0.1211
20	44.85	44.9267	44.9334	45.2411	45.0337	-0.1837
21	44.93	44.9302	44.9231	44.8762	44.9098	0.0202
22	44.60	44.9337	44.6042	44.9180	44.8186	-0.2186
23	44.83	44.9372	44.6870	44.4097	44.6780	0.1520
24	44.75	44.9407	44.6597	44.6539	44.7514	-0.0014
25	45.05	44.9442	44.9345	44.5801	44.8196	0.2304
26	45.14	44.9477	45.1430	45.0291	45.0400	0.1000
27	44.87	44.9512	44.9574	45.2656	45.0581	-0.1881
28	45.04	44.9547	45.0298	44.8937	44.9594	0.0806
29	45.24	44.9582	45.2641	45.0744	45.0989	0.1411
30	45.25	44.9617	45.3721	45.4234	45.2524	-0.0024

FONTE: O autor (2010)

4.3.2.1 Previsões e Intervalo “Bootstrap” para PT32

A tabela 18, adiante, mostra os seis valores da série guardados para teste de desempenho, a previsões para $L_h(\Theta = 0)^{RL}$, $L_h(\Theta = 1.6)^{AELB}$ e $L_h(\Theta = 2)^{AES}$, a previsão para o Método Theta $\hat{Y}_t(h)$, a avaliação da qualidade da previsão de acordo com o critério MAPE, EQM e RMSE.

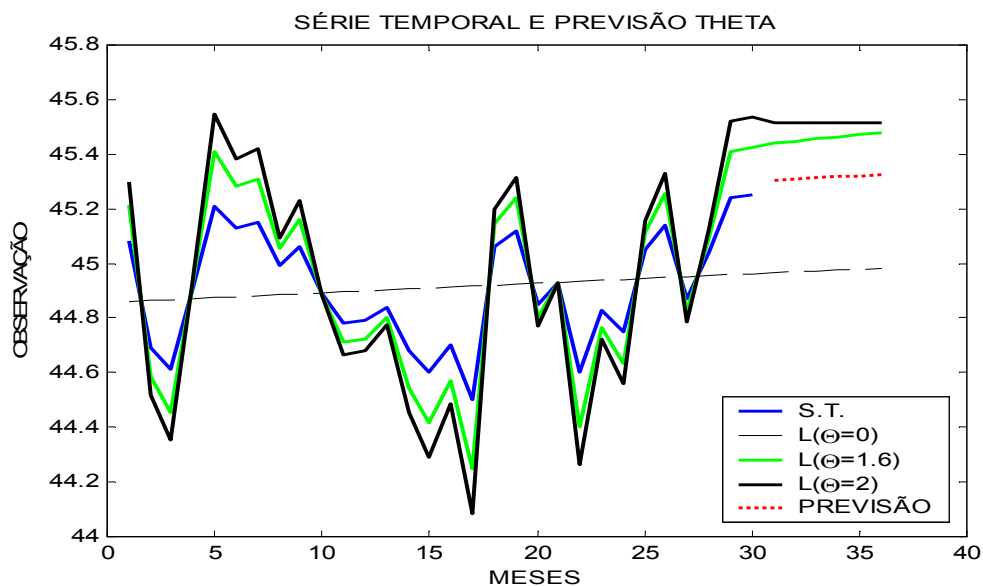
TABELA 18 - PERÍODO, VALOR REAL, LINHAS DE REGRESSÃO E ALISAMENTO, PREVISÕES THETA, ERRO MÉDIO PERCENTUAL ABSOLUTO (MAPE) E ERRO QUADRÁTICO MÉDIO (EQM)

Per. (h)	Obs. (teste)	$\underline{L}_h(\Theta = 0)^{RL}$	$\underline{L}_h(\Theta = 1.6)^{AELB}$	$\underline{L}_h(\Theta = 2)^{AES}$	$\hat{Y}_t(h)$	MAPE	EQM
1	44.93	44.9652	45.4407	45.5130	45.306	0.8375	0.141
2	45.21	44.9687	45.4483	45.5130	45.310	0.2212	0.010
3	45.10	44.9722	45.4559	45.5130	45.313	0.4738	0.045
4	45.18	44.9757	45.4635	45.5130	45.317	0.3041	0.018
5	45.09	44.9792	45.4711	45.5130	45.321	0.5125	0.053
6	45.15	44.9826	45.4787	45.5130	45.324	0.3872	0.030
Méd. RMSE						0.456%	0.050 0.2236

FONTE: O autor (2010)

O RMSE para a série decomposta em três linhas e pela extrapolação por tendência linear, alisamento exponencial linear de Brown e alisamento exponencial simples é igual a 0.2236. A medida de desempenho médio de acordo com o critério MAPE é de 0.456%. O gráfico 12, adiante, mostra a série temporal analisada, a linha de regressão linear, de alisamento exponencial e as previsões para o método theta.

GRÁFICO 12 - SÉRIE TEMPORAL, $L(\Theta=0)^{RL}$, $L(\Theta=1.6)^{AELB}$, $L(\Theta=2)^{AES}$ E PREVISÃO THETA



FONTE: O autor (2010)

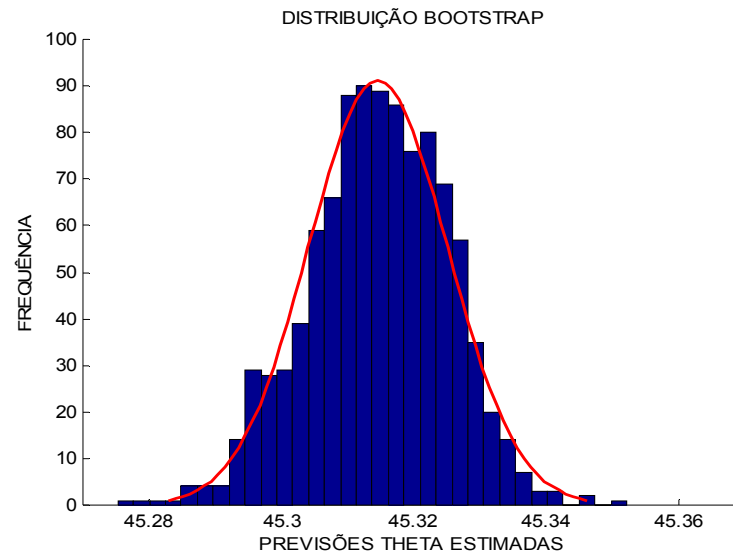
A tabela 19, adiante, mostra os valores observados, as previsões para o Método Theta, Desvio Padrão “bootstrap”, EQM “bootstrap” e os Limites Inferiores e Superiores de Confiança.

TABELA 19 - VALOR REAL DA SÉRIE, RESULTADO DA PREVISÃO THETA, DESVIO PADRÃO “BOOTSTRAP”, ERRO QUADRÁTICO MÉDIO “BOOTSTRAP” E O INTERVALO DE CONFIANÇA (NÍVEL DE 5% E 1000 REPLICAÇÕES)

Período (h)	Observado $\underline{Y}_t(h)$	P.Theta $\hat{\underline{Y}}_t(h)$	Desvio- Padrão “bootstrap”	EQM “bootstrap”	Limite Inf. 95%	Limite Sup. 95%
1	44.9300	45.3063	0.0065	0.51×10^{-4}	45.2932	45.3338
2	45.2100	45.3100	0.0072	0.67×10^{-4}	45.2999	45.3269
3	45.1000	45.3137	0.0076	0.87×10^{-4}	45.3030	45.3329
4	45.1800	45.3174	0.0083	1.08×10^{-4}	45.3061	45.3385
5	45.0900	45.3211	0.0096	1.42×10^{-4}	45.3084	45.3453
6	45.1500	45.3248	0.0105	1.78×10^{-4}	45.3115	45.3520
				1.06×10^{-4}		

FONTE: O autor (2010)

Analisando a tabela 19, nota-se que o EQM “bootstrap” relativo as previsões obtidas, aparecem em ordem crescente, significando que a cada horizonte de previsão os limites de confiança aparecem em um intervalo mais amplo. Tornando as previsões menos confiáveis a medida que o horizonte de previsão aumenta. O intervalo “Bootstrap” para a previsão no horizonte de h períodos à frente para a série temporal, aplicou-se o Método “Bootstrap” com um número de replicações $B = 1000$. O histograma das estimativas “bootstrap” da previsão $\hat{Y}_t^*(h)$, para a primeira previsão, fig. 10 a seguir, mostra a distribuição “bootstrap” para $\hat{Y}_t^*(1)$. No gráfico percebe-se um nível alto de simetria, o que sugere um modelo Gaussiano para esses valores obtidos.

FIGURA 10 - DISTRIBUIÇÃO “BOOTSTRAP” DE PT32 PARA $\hat{Y}_T^*(1)$ 

FONTE: O autor (2010)

4.4 COMPARAÇÃO ENTRE AS PREVISÕES OBTIDAS PELO MÉTODO THETA COM OUTROS MÉTODOS

As tabelas 20 e 21, a seguir, mostram os resultados das previsões obtidas por cada método aplicado neste trabalho e seus respectivos RMSE e erros MAPE. Apresentando também as previsões obtidas pelo Software Statgraphics utilizando métodos tradicionais otimizados pelo menor RMSE.

TABELA 20 - PREVISÕES E MAPES's OBTIDOS PELO MÉTODO THETA DECOMPOSTO EM DUAS LINHAS THETA

H	$\Theta=0, \Theta=2$ RL e AES RMSE=0.1606		$\Theta=0, \Theta=2$ RL e AELB RMSE=0.1897		$\Theta=0, \Theta=2$ RL e AEH RMSE=0.2964	
	Previsão	MAPE	Previsão	MAPE	Previsão	MAPE
1	45.2391	0.6880	45.2628	0.7406%	45.2902	0.8017%
2	45.2409	0.0682	45.2692	0.1310%	45.3315	0.2688%
3	45.2426	0.3162	45.2757	0.3896%	45.3728	0.6049%
4	45.2443	0.1424	45.2822	0.2262%	45.4141	0.5182%
5	45.2461	0.3462	45.2887	0.4406%	45.4554	0.8104%
6	45.2478	0.2167	45.2951	0.3215%	45.4967	0.7679%
	MAPE médio	0.2963%		0.3742%		0.6287%

FONTE: O autor (2010)

TABELA 21 - PREVISÕES E MAPE's OBTIDOS PELO MÉTODO THETA DECOMPOSTO EM TRÊS LINHAS THETA E PREVISÕES E MAPE OBTIDOS POR OUTRO MÉTODO DE PREVISÃO AJUSTADO PELO MENOR RMSE.

h	$\Theta=0, \Theta=0,5$ e $\Theta=2$ RL, AES e AELB RMSE=0.1378		$\Theta=0, \Theta=1,6$ e $\Theta=2$ RL, AELB e AES RMSE=0.2236		ARMA(1,0) RMSE=0.1897	
	Previsão	MAPE	Previsão	MAPE	Previsão	MAPE
1	45.2083	0.6194%	45.3063	0.8375%	45.0868	0,3489%
2	45.2126	0.0058%	45.3100	0.2212%	45,0054	0,4525%
3	45.2169	0.2593%	45.3137	0.4738%	44.9648	0,2997%
4	45.2212	0.0913%	45.3174	0.3041%	44.9445	0,5212%
5	45.2256	0.3006%	45.3211	0.5125%	44.9344	0,3450%
6	45.2299	0.1769%	45.3248	0.3872%	44.9293	0,4888%
MAPE médio		0.2422%				0.4095%

FONTE: O autor (2010)

A tabela 20 permite afirmar que o método theta em sua aplicação mais simples $L(\Theta=0)$ e $L(\Theta=2)$, obteve seu melhor desempenho utilizando o método de alisamento exponencial simples (AES) para extrapolação da $L(\Theta=2)$, obtendo um erro médio percentual absoluto de 0,2963%. Já para a decomposição em três linhas theta, como ilustra a tabela 21 o melhor resultado foi alcançado pelas linhas theta $L(\Theta=0)$, $L(\Theta=0.5)$ e $L(\Theta=2)$ extrapoladas por Tendência Linear, Alisamento Exponencial Simples e Alisamento Exponencial Linear de Brown, respectivamente, obtendo um erro médio percentual absoluto de 0,2963%. Esta combinação de três linhas theta também foi a que obteve os melhores resultados para o método e também o melhor resultado se comparado ao alcançado pelo método ajustado pela menor Raiz do Erro Quadrático Médio (RMSE) obtido pelo programa Statgraphics, que obteve RMSE igual a 0.1897 e MAPE igual a 0.4095.

5 CONCLUSÃO E SUGESTÕES PARA TRABALHOS FUTUROS

O objetivo fundamental de construir intervalos de confiança para as previsões pelo método Theta foi alcançado pela técnica de computação intensiva “Bootstrap”. O método Theta também mostrou ser uma boa alternativa para a previsão de séries temporais, pois os resultados obtidos apontam uma vantagem significativa sobre métodos convencionais de previsão.

O método é simples, não requer um grande treinamento e bases estatísticas. A partir dos resultados obtidos, nota-se que a maior complexidade de um modelo não resulta necessariamente em melhor resultado na modelagem dos dados. Na sua aplicação mais simples, pela decomposição da série em regressão e alisamento exponencial simples obteve-se resultados no mínimo equivalentes aos demais métodos automáticos. Os resultados dos erros MAPE's confirmam os resultados favoráveis ao método da competição M3 de Makridakis e Hibon (2000). Para a construção dos intervalos de confiança foi utilizado o método de computação intensiva “Bootstrap”, pelo qual se obteve os intervalos percentílicos de 95 % de confiança. Os dados da distribuição “bootstrap” mostraram comportamento simétrico, o que sugere um modelo de estimativas normalmente distribuídas com 1000 replicações. O problema para determinar quantas replicações B são necessárias para a obtenção de boas estimativas dos limites inferiores e superiores de intervalos de confiança pelo método “bootstrap” é discutido em Efron & Tibshirani (1986).

Como sugestão para novos trabalhos, podem ser citadas:

- A decomposição da série em quatro ou mais linhas theta.
- Aplicar outras técnicas de alisamento tais como Alisamento Exponencial Quadrático de Brown e Alisamento Exponencial de Holt Winters.
- Analisar séries reais tais quais apresentem padrões diferentes e tamanhos de amostra maiores.

REFERÊNCIAS

- ASSIMAKOPOULOS, V.; NIKOLOPOULOS, K. **The theta model: a decomposition approach to forecasting**. International Journal of Forecasting v.16 p. 521 –530, 2000.
- BARROS, Mônica. **Processos Estocásticos**. Editora, 2003 Disponível em: [http://www.mbarros.com/sitebuildercontent/sitebuilderfiles/Series Temporais.pdf](http://www.mbarros.com/sitebuildercontent/sitebuilderfiles/Series_Temporais.pdf); Último acesso em março de 2006.
- BOX, G. E. P.; JENKINS, G. M. & GWILYM M. **Time Series Analysis: forecasting and control**. San Francisco, Holden-Day, 1976.
- BOX, G. E. P.; JENKINS, G. M. & REINSEL, G. C. – **Time Series Analysis: forecasting and control**, 3.rd ed., 1994, Prentice Hall, New Jersey, 587p.
- BRESSAN, A. A.; LIMA, J. E., **Modelos de previsão de preços aplicados aos contratos futuros de boi gordo na BM&F**. Revista Nova Economia, Belo Horizonte, v.12, p.117-140, 2002.
- BROWN, R.G.; MEYER, The Fundamental Theorem of Exponential Smoothing. Operations Research, 9, nr.5, pp. 673-85 1961.
- CHAVES NETO, A., **“Bootstrap” em Séries Temporais**. (Tese de Doutorado, 1991). Pontifícia Universidade Católica do Rio de Janeiro. PUC-RJ.
- CHAVES NETO, A. **Análise de Séries Temporais**. Notas de aula. Departamento de Estatística, Universidade Federal do Paraná, Curitiba, 2004.
- CORRÊA, H. L.; GIANESI, I. G. N.; CAON, M. **Planejamento, programação e controle da produção: MRP II/EPR: conceitos, uso e implantação**. 3^a ed. São Paulo: Giancesi Corrêa & Associados: Atlas, 2000.
- DAVISON, A. C.; HINKLEY, D. V., **“Bootstrap” Methods and their Application**. Cambridge University Press, 1997.
- EFRON, B. **“Bootstrap” Methods: another look at the jackknife**. The Annals of Statistics 1.:1-20, 1979.
- EFRON, B. **The Jackknife, the “Bootstrap”, and Other Resampling Plans**. Number 38 in CBMS-NSF Regional Conference Series in Applied Mathematics. 1982. Philadelphia: SIAM.
- EFRON, B. and TIBSHIRANI, R. J. **An introduction to the “bootstrap”**, volume 57 of Monographs on Statistics and Applied Probability. Chapman and Hall, New York, 1993.

EHLERS, Ricardo S. **Análise de Séries Temporais**. Departamento de Estatística, UFPR, 2005. Disponível em <http://leg.est.ufpr.br/~ehlers/notas>. Acessado em: 06/09/2009.

ELSAYED, E. A. & BOUCHER, T. O. **Analysis and control of production systems.**, 2.nd ed., 1994, Prentice Hall, New Jersey, 450 p.

GOOIJER, J.G.; HYNDMAN R.J. **25 years of time series forecasting**. International Journal of Forecasting, Vol 22, p. 443–473, 2006.

GRANGER, C. W. J. & NEWBOLD, P.; **Forecasting economic time series**. 1.st ed., 1986, Academic Press, New York.

HESTERBERG, T.; MOORE, D. S.; MONAGHAN, S.; CLIPSON, A.; EPSTEIN, R.; **“Bootstrap” methods and permutation tests**, In: The practice of business statistics. New York: W. H. Freeman, 2003.

HYNDMAN, R.J.;BILLAH, B. **Unmasking the Theta method**. International Journal of Forecasting, v.19 (2), p. 287-290, 2003.

KRAJEWSKI, L. J. & RITZMAN, L. P. **Operations management, strategy and analysis**, 5^a ed., 1998, Addison-Wesley, Reading, MA, 880p.

LIEBEL, M. J., **Previsão de Receitas Tributárias – O caso do ICMS no Estado do Paraná**. 2004. Dissertação (Mestrado Profissionalizante em Engenharia) – Universidade Federal do Rio Grande do Sul – RS.

MAKRIDAKIS, S.; HIBON, M. **The M3-Competition: results, conclusions and implications**. International Journal of Forecasting, v. 16, p. 451 –476, 2000.

MAKRIDAKIS, S. G.; WHEELWRIGHT, S. C; HYDMAN, R. J. **Forecasting: methods and applications**. 3^a ed. New York: John Wiley & Sons, Inc. 1998. 642p.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; McGEE, V. E. **Forecasting: Methods and Applications**. 2^a ed.. Nova York: John Wile & Sons, 1983.

MARTINEZ-ESPINOSA, M.; SANDANIELO, V. L. M.; LOUZADA-NETO, F.; **O método de “bootstrap” para o estudo de dados de fadiga dos materiais**. Ver. Mat. Estat., São Paulo, v.24, n.2, p.37-50, 2006.

MILLS, T. C.; **Time series techniques for economist**. 1990, Cambridge University Press, Cambridge.

MONTGOMERY, D. C. – **Introduction to statistical quality control**, 3.rd ed. 1997, John Wiley & Sons, Inc., New York, 677p.

MORETTIN, P. A. and TOLOI, C. M. C. (2006). **Análise de Séries Temporais**. BLÜCHER.

MORETTIN, P. A.; TOLOI, C. M. de C. **Modelos para Previsão de Séries Temporais**. In: Colóquio Brasileiro de Matemática, 13º, 1981, Poços de Caldas-MG, IMPA, Rio de Janeiro, 1981.

NIKOPOULOS K., & THOMAKOS, D., & PETROPOULOS, F & ASSIMAKOPOULOS, V., 2009. **"Theta Model Forecasts for Financial Time Series: A Case Study in the S&P500,"** Working Papers 0033, University of Peloponnese, Department of Economics.

NIKOLOPOULOS, K.; ASSIMAKOPOULOS, V. **Fathoming the Theta model**. 25th International Symposium on Forecasting, ISF, San Antonio, Texas, USA, 2005.

NIKOLOPOULOS, K.; ASSIMAKOPOULOS, V. **Advances in the theta model**. University of Peloponnese, Department of Economics, 2008.

PELEGRINI, F. R. & FOGLIATTO F. S. – **Metodologia para Implantação de Sistemas de Previsão de Demanda**. Universidade Federal do Rio Grande do Sul (UFRGS), 2002.

QUENOUILLE, M. H. **Approximate tests of correlation in time-series**. Journal of Statistical Computation and Simulation, 8:75-80, 1949.

SOUZA, Reinaldo Castro. **Modelos Estruturais para Previsão de Séries Temporais : Abordagens Clássica e Bayesiana**. In : 17º Colóquio Brasileiro de Matemática. Rio de Janeiro, 1989.

STATSOFT, INC. **Electronic Statistic Textbook**. Tulsa, OK: Statsoft, 1999. Web: <http://www.statsoft.com/textbook/stathome/html>.

STEVENSON, W. J. – **Estatística Aplicada à Administração**. São Paulo. Editora Harbra Ltda., 1981, 495p.

TRENTIN, M. G. – **Planejamento Estratégico de Materiais: Uma Aplicação dos Métodos Automáticos de Previsão de Séries Temporais**, 2002. Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Universidade Federal do Paraná – PR.

VICENTE, J. R. **Modelos Estruturais para previsão das produções brasileiras de carne de frango e ovos**. Revista de Economia e Sociologia Rural, Brasília, v.30, n.4, p.305-319, 1992.

WHEELWRIGHT, Steven C.; MAKRIDAKIS, Spyros. **Forecasting Methods for Management**. 4th edition. New York : John Wiley & Sons Inc, 1985

WOLD, H. O. **A Study in the Analysis of Stationary Time Series**. Almqvist and Wiksell, Uppsala, 1938.

APÊNDICES

APÊNDICE A - PROGRAMA PARA A PREVISÃO DE SÉRIE TEMPORAL ATRAVÉS DO MÉTODO THETA UTILIZANDO O SOFTWARE MATLAB

```
function[out]=prevTheta(Z)

% Z:dados originais;
% L(0): regressao linear dos dados (tetha=0)
format short
Z=Z; %entrar com os valores da série;
[t,c]=size(Z); %valor da variável x;
disp('*****')
disp(' * quantidades de linhas theta *')
disp('*****')
disp(' * T = 2, L(0) e L(2) *')
disp(' * T = 3, L(0), L(?) e L(2) *')
disp('*****')
disp(' ')
disp(' ')
T=input(' ENTRAR COM A OPÇÃO T=');
if T==2

disp('*****')
disp(' * MÉTODO DE ALISAMENTO EXPONENCIAL *')
disp('*****')
disp(' * M = 1, A.E.SIMPLES *')
disp(' * M = 2, A.E.BROWN *')
disp(' * M = 3, A.E.HOLT *')
disp('*****')
M=input(' ENTRAR COM A OPÇÃO M=');

if M==1

n=input(' ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS n=');

Zteste=Z(c-n+1:c);
c=c-n;
Z=Z(1:c);
x=1:c;
X=[(ones(c,1)) x']; %matriz X de ordem nxp
Beta=(inv(X'*X))*X'*Z';
L0=X*Beta;
L2=2*Z'-L0; %combinação theta(2);

Comb=(L0+L2)/2;
disp('série original L(Theta=0) L(Theta=2) combinação')
[ Z' L0 L2 Comb ]
Zteste
As=AEScoluna(L2');
Asimples=As(1:c,1)
```

```

PAES=AES(L2');
Pcomb=(X*Beta+Asimples)/2      %vetor de residuos
Res=Z'-Pcomb
Beta1=Beta;

for k=1:n
    h=[1 c+k];
    PAES;
    YlinearP=h*Beta1;
    PrevisaoX=(YlinearP+PAES)/2;
    Zteste(k,1)
    MAPE(k)=(abs((Zteste(k,1)-PrevisaoX))/Zteste(k,1))*100;
    EQM(k)=(PrevisaoX-Zteste(k,1))^2;

for i=1:1000
    resb=(bootrsp(Res,1));
    combin=Pcomb+resb;
    Beta=(inv(X'*X))*X'*combin;
    L00=X*Beta;          %amostra de pseudo-dados (somado com vetor bootstrap de
residuos)
    h=[1 c+k];          %h para 1º periodo a frente;
    Ylinear=h*Beta;
    L22=2*Z'-L00;
    PrevAES=AES(L22');
    Asimp=AEScoluna(L22');
    Asimpl=Asimp(1:c,1);
    Pcomb=(L00+Asimpl)/2;
    PrevTheta=(PrevAES+Ylinear)/2;
    Yy1(i)=PrevTheta;
    EQM1(i)=(PrevTheta-PrevisaoX)^2;    %ERRO QUADRATICO

end
EST1=full(Yy1);          %estimativas para (RL+AES)/2 no 1º periodo;
dpadraoBOOT(k)=std(EST1);
EQMboot(k)=mean(full(EQM1));

alpa=0.05;
INT=sort(EST1');          %ordenação dos valores da amostra bootstrap em ordem crescente;
q1=1000*alpa/2;
q2=1000-q1+1;
lo1=INT(1,q1);
up1=INT(1,q2);

AL(k)=PAES;
YLin(k)=YlinearP;
inf(k)=lo1;
sup(k)=up1;
pthet(k)=PrevisaoX;

end

LTheta0=full(YLin)';
LTheta2=full(AL)';

```

```

Previsao=full(pthet);
intinf=full(inf);
intsup=full(sup);
MAPES=full(MAPE);
EQMs=full(EQM);
stPADRAO=full(dpadroBOOT);
EQMbootstrap=full(EQMboot);
disp(' periodo L(Theta=0) L(Theta=2) Previsao inf sup')

[(1:n)' LTheta0 LTheta2 Previsao intinf intsup ]

disp(' MAPE EQM STD EQMboot')
[MAPES EQMs stPADRAO EQMbootstrap]
MAPEmedio=mean(MAPES)
EQMmedio=mean(EQMs)
EQMbootmedio=mean(EQMbootstrap)

%GRAFICOS;

xl=[1:c];
xll=[1:c+n];
xp=[c+1:c+n];
plot(xl,Z,'*-'), legend('S.T.'), title('SÉRIE TEMPORAL'),
Xlabel('MESES'),Ylabel('OBSERVAÇÃO')
figure
plot(xl,Z,'b-',xll,[L0;LTheta0'],'k--',xll,[L2;LTheta2'],'k-',xp,Previsao,'r:')
legend('S.T.','L(\Theta=0)','L(\Theta=2)','PREVISÃO')
title('SÉRIE TEMPORAL E PREVISÃO THETA'),Xlabel('MESES'),Ylabel('OBSERVAÇÃO')

%%

if k==1
figure
histfit(EST1)
box off
xlabel('PREVISÕES THETA ESTIMADAS') % legenda do eixo horizontal
ylabel('FREQUÊNCIA') %legenda do eixo vertical
title('DISTRIBUIÇÃO BOOTSTRAP')
end

elseif M==2

n=input(' ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS n=');

Zteste=Z(c-n+1:c);
c=c-n;
Z=Z(1:c);
x=1:c;
X=[(ones(c,1)) x']; %matriz X de ordem nxp
Beta=(inv(X'*X))*X'*Z';
L0=X*Beta;
L2=2*Z'-L0; %combinação theta(2);

```

```

Comb=(L0+L2)/2;
disp('série original L(Theta=1) L(Theta=2) combinação')
[      Z'      L0      L2      Comb ]
Zteste
Ab=BROWNcoluna(L2');
Abrown=Ab
Pcomb=(X*Beta+Abrown)/2      %vetor de residuos
Res=Z'-Pcomb
Beta1=Beta;

for k=1:n
    h=[1 c+k];
    PBROWN=BROWN(L2');
    PBROW=PBROWN(k,1);
    YlinearP=h*Beta1;
    PrevisaoX=(YlinearP+PBROW)/2;
    Zteste(k,1)
    MAPE(k)=(abs((Zteste(k,1)-PrevisaoX))/Zteste(k,1))*100;
    EQM(k)=(PrevisaoX-Zteste(k,1))^2;

for i=1:1000
    resb=(bootrsp(Res,1));
    combin=Pcomb+resb;
    Beta=(inv(X'*X))*X'*combin;
    L00=X*Beta;      %amostra de pseudo-dados (somado com vetor bootstrap de
residuos)
    h=[1 c+k];      %h para 1º periodo a frente;
    Ylinear=h*Beta;
    L22=2*Z'-L00;
    Abrow=BROWNcoluna(L22');
    Abrown=Abrow(1:c,1);
    Pcomb=(L00+Abrown)/2;
    PrevBRO=BROWN(L22');
    PrevBROWN=PrevBRO(k,1);
    PrevTheta=(PrevBROWN+Ylinear)/2;
    Yy1(i)=PrevTheta;
    EQM1(i)=(PrevTheta-PrevisaoX)^2;

end
EST1=full(Yy1)';      %estimativas para (RL+AES)/2 no 1º periodo;
dpadiaoBOOT(k)=std(EST1);
EQMboot(k)=mean(full(EQM1));

alpa=0.05;
INT=sort(EST1)';      %ordenação dos valores da amostra bootstrap em ordem crescente;
q1=1000*alpa/2;
q2=1000-q1+1;
lo1=INT(1,q1);
up1=INT(1,q2);

AB(k)=PBROW;
YLin(k)=YlinearP;

```

```

inf(k)=lo1;
sup(k)=up1;
pthet(k)=PrevisaoX;

end

LTheta0=full(YLin)';
LTheta2=full(AB)';
Previsao=full(pthet)';
intinf=full(inf)';
intsup=full(sup)';
MAPES=full(MAPE)';
EQMs=full(EQM)';
stPADRAO=full(dpadraoBOOT)';
EQMbootstrap=full(EQMboot)';

disp(' periodo L(Theta=0) L(Theta=2) Previsao inf sup')

[(1:n)' LTheta0 LTheta2 Previsao intinf intsup ]

disp(' MAPE EQM STD EQMboot')
[MAPES EQMs stPADRAO EQMbootstrap]
MAPEmedio=mean(MAPES)
EQMmedio=mean(EQMs)
EQMbootmedio=mean(EQMbootstrap)

%GRAFICOS;

xl=[1:c];
xll=[1:c+n];
xp=[c+1:c+n];

plot(xl,Z), legend('S.T.'), title('Série Temporal')

plot(xl,Z,'b-',xll,[L0;LTheta0]','k--',xll,[L2;LTheta2]','k-',xp,Previsao','r:')
legend('S.T.','L(\Theta=0)','L(\Theta=2)','Previsão')
title('SÉRIE TEMPORAL E PREVISÃO THETA'),Xlabel('MESES'),Ylabel('OBSERVAÇÃO')
%%

if k==1
figure
histfit(EST1)
box off
xlabel('PREVISÕES THETA ESTIMADAS') % legenda do eixo horizontal
ylabel('FREQUÊNCIA') %legenda do eixo vertical
title('DISTRIBUIÇÃO BOOTSTRAP')
end

elseif M==3

n=input(' ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS n=');

Zteste=Z(c-n+1:c);

```

```

c=c-n;
Z=Z(1:c);
x=1:c;
X=[(ones(c,1)) x'];      %matriz X de ordem nxp
Beta=(inv(X'*X))*X'*Z';
L0=X*Beta;
L2=2*Z'-L0;              %combinação theta(2);

Comb=(L0+L2)/2;
disp('série original L(Theta=1) L(Theta=2) combinação')
[      Z'      L0      L2      Comb ]
Zteste
Ah=HOLTcoluna(L2');
Aholt=Ah
Pcomb=(X*Beta+Aholt)/2    %vetor de residuos
Res=Z'-Pcomb
Beta1=Beta;

for k=1:n
    h=[1 c+k];
    PHOLT=HOLT(L2');
    PHOL=PHOLT(k,1);
    YlinearP=h*Beta1;
    PrevisaoX=(YlinearP+PHOL)/2;
    Zteste(k,1)
    MAPE(k)=(abs((Zteste(k,1)-PrevisaoX))/Zteste(k,1))*100;
    EQM(k)=(PrevisaoX-Zteste(k,1))^2;

for i=1:1000
    resb=(bootrsp(Res,1));
    combin=Pcomb+resb;
    Beta3=(inv(X'*X))*X'*combin;
    L00=X*Beta3;          %amostra de pseudo-dados (somado com vetor bootstrap de
residuos)
    h=[1 c+k];           %h para 1º periodo a frente;
    Ylinear3=h*Beta;
    L22=2*Z'-L00;
    Ahol=HOLTcoluna(L22');
    Aholt=Ahol(1:c,1);
    Pcomb=(L00+Aholt)/2;
    PrevHOL=HOLT(L22');
    PrevHOLT=PrevHOL(k,1);
    PrevTheta3=(PrevHOLT+Ylinear3)/2;
    Yy3(i)=PrevTheta3;
    EQM3(i)=(PrevTheta3-PrevisaoX)^2;

end
EST3=full(Yy3);          %estimativas para (RL+AES)/2 no 1º periodo;
dpadiaoBOOT(k)=std(EST3);
EQMboot(k)=mean(full(EQM3));

alpa=0.05;
INT=sort(EST3);          %ordenação dos valores da amostra bootstrap em ordem crescente;

```

```

q1=1000*alpha/2;
q2=1000-q1+1;
lo1=INT(1,q1);
up1=INT(1,q2);

AHO(k)=PHOL;
YLin(k)=YlinearP;
inf(k)=lo1;
sup(k)=up1;
pthet(k)=PrevisaoX;

end

LTheta0=full(YLin);
LTheta2=full(AHO);
Previsao=full(pthet);
intinf=full(inf);
intsup=full(sup);
MAPES=full(MAPE);
EQMs=full(EQM);
stPADRAO=full(dpadraoBOOT);
EQMbootstrap=full(EQMboot);

disp(' periodo L(Theta=0) L(Theta=2) Previsao inf sup')

[(1:n)' LTheta0 LTheta2 Previsao intinf intsup ]

disp(' MAPE EQM STD EQMboot')
[MAPES EQMs stPADRAO EQMbootstrap]
MAPEmedio=mean(MAPES)
EQMmedio=mean(EQMs)
EQMbootmedio=mean(EQMbootstrap)

%GRAFICOS;

xl=[1:c];
xll=[1:c+n];
xp=[c+1:c+n];

plot(xl,Z), legend('S.T.'), title('Série Temporal')

plot(xl,Z,'b-',xll,[L0;LTheta0'],'k--',xll,[L2;LTheta2'],'k-',xp,Previsao,'r:')
legend('S.T.','L(\Theta=0)','L(\Theta=2)','PREVISÃO')
title('SÉRIE TEMPORAL E PREVISÃO THETA'),Xlabel('MESES'),Ylabel('OBSERVAÇÃO')
%%

if k==1
figure
histfit(EST3)
box off
xlabel('PREVISÕES THETA ESTIMADAS') % legenda do eixo horizontal
ylabel('FREQUÊNCIA') %legenda do eixo vertical
title('DISTRIBUIÇÃO BOOTSTRAP')

```

```
end
end
```

```
elseif T==3
    ORD=input('    ENTRAR COM A ORDEM DE AE, SE 1 = AELB e AES, SE 2 = AES E AELB ORD=');
    if ORD==1
        v=input('    ENTRAR COM O VALOR DO COEFICIENTE THETA v=');
        n=input('    ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS n=');
```

```
Zteste=Z(c-n+1:c)';
c=c-n;
Z=Z(1:c);
x=1:c;
X=[(ones(c,1)) x'];          %matriz X de ordem nxp
Beta=(inv(X'*X))*X'*Z';
L0=X*Beta;
L05=v*Z'+(1-v)*L0;
L2=2*Z'-L0;                  %combinação theta(2);
```

```
Comb=(L0+L05+L2)/3;
disp('série original L(Theta=1) L(Theta=00) L(Theta=2) combinação')
[    Z'      L0      L05      L2      Comb ]
Zteste
As=AEScoluna(L2');
Asimples=As(1:c,1)
PAES=AES(L2');
Ab=BROWNcoluna(L05');
Abrown=Ab(1:c,1)
Pcomb=(X*Beta+Asimples+Abrown)/3      %vetor de residuos
Res=Z'-Pcomb
Beta1=Beta;
```

```
for k=1:n
    h=[1 c+k];
    PAES;
    YlinearP=h*Beta1;
    PBROWN=BROWN(L05');
    PBROW=PBROWN(k,1);
    PrevisaoX1=(YlinearP+PAES+PBROW)/3;
    Zteste(k,1)
    MAPE(k)=(abs((Zteste(k,1)-PrevisaoX1))/Zteste(k,1))*100;
    EQM(k)=(PrevisaoX1-Zteste(k,1))^2;
```

```
for i=1:1000
    resb=(bootrsp(Res,1));
    combin=Pcomb+resb;
    Beta=(inv(X'*X))*X'*combin;
    L00=X*Beta;              %amostra de pseudo-dados (somado com vetor bootstrap de
    residuos)
    h=[1 c+k];              %h para 1º periodo a frente;
```



```

Ylinear=h*Beta;
L22=2*Z'-L00;
PrevAES=AES(L22');
L050=v*Z'+(1-v)*L00;
PrevBROW=BROWN(L050');
PrevBROWN=PrevBROW(k,1);
Asim=AEScoluna(L22');
Asimples=Asim(1:c,1);
Abro=BROWNcoluna(L050');
Abrow=Abro(1:c,1);
Pcomb=(L00+Asimples+Abrow)/3;
PrevTheta=(PrevAES+Ylinear+PrevBROWN)/3;
Yy1(i)=PrevTheta;
EQM1(i)=(PrevTheta-PrevisaoX1)^2;

end
EST1=full(Yy1);      %estimativas para (RL+AES)/2 no 1º periodo;
dpadraoBOOT(k)=std(EST1);
EQMboot(k)=mean(full(EQM1));

alpa=0.05;
INT=sort(EST1);      %ordenação dos valores da amostra bootstrap em ordem crescente;
q1=1000*alpa/2;
q2=1000-q1+1;
lo1=INT(1,q1);
up1=INT(1,q2);

AL(k)=PAES;
YLin(k)=YlinearP;
AB(k)=PBROW;
inf(k)=lo1;
sup(k)=up1;
pthet(k)=PrevisaoX1;

end

LTheta0=full(YLin);
LTheta05=full(AB);
LTheta2=full(AL);
Previsao=full(pthet);
intinf=full(inf);
intsup=full(sup);
MAPES=full(MAPE);
EQMs=full(EQM);
stPADRAO=full(dpadraoBOOT);
EQMbootstrap=full(EQMboot);

disp(' periodo L(Theta=0) L(Theta=0.5) L(Theta=2) Previsao inf sup')

[(1:n)' LTheta0 LTheta05 LTheta2 Previsao intinf intsup ]

disp(' MAPE EQM STD EQMboot')
[MAPES EQMs stPADRAO EQMbootstrap]
MAPEmedio=mean(MAPES)

```

```
EQMmedio=mean(EQMs)
EQMbootmedio=mean(EQMbootstrap)
```

```
%GRAFICOS;
```

```
xl=[1:c];
xll=[1:c+n];
xp=[c+1:c+n];
```

```
plot(xl,Z), legend('S.T.'), title('Série Temporal')
```

```
plot(xl,Z,'b-',xll,[L0;LTheta0'],'k--',xll,[L05;LTheta05'],'g-',xll,[L2;LTheta2'],'k-',xp,Previsao','r:')
legend('S.T.','L(\Theta=0)','L(\Theta=1.6)','L(\Theta=2)','PREVISÃO')
title('SÉRIE TEMPORAL E PREVISÃO THETA'),Xlabel('MESES'),Ylabel('OBSERVAÇÃO')
%%
```

```
if k==1
    figure
    histfit(EST1)
    box off
    xlabel('PREVISÕES THETA ESTIMADAS') % legenda do eixo horizontal
    ylabel('FREQUÊNCIA') %legenda do eixo vertical
    title('DISTRIBUIÇÃO BOOTSTRAP')
end
```

```
elseif ORD==2
```

```
v=input(' ENTRAR COM O VALOR DO COEFICIENTE THETA v=');
n=input(' ENTRAR COM O NÚMERO DE PREVISÕES A SEREM REALIZADAS n=');
```

```
Zteste=Z(c-n+1:c);
c=c-n;
Z=Z(1:c);
x=1:c;
X=[(ones(c,1)) x']; %matriz X de ordem nxp
Beta=(inv(X'*X))*X'*Z';
L0=X*Beta;
L05=v*Z'+(1-v)*L0;
L2=2*Z'-L0; %combinação theta(2);
```

```
Comb=(L0+L05+L2)/3;
disp('série original L(Theta=1) L(Theta=00) L(Theta=2) combinação')
[ Z' L0 L05 L2 Comb ]
Zteste
Ab=BROWNcoluna(L2');
Abrown=Ab(1:c,1)
PAES=AES(L05');
As=AEScoluna(L05');
Asimples=As(1:c,1)
Pcomb=(X*Beta+Asimples+Abrown)/3 %vetor de residuos
Res=Z'-Pcomb
```

```

Beta1=Beta;

for k=1:n
    h=[1 c+k];
    PAES;
    YlinearP=h*Beta1;
    PBROWN=BROWN(L2');
    PBROW=PBROWN(k,1);
    PrevisaoX=(YlinearP+PAES+PBROW)/3;
    Zteste(k,1)
    MAPE(k)=(abs((Zteste(k,1)-PrevisaoX))/Zteste(k,1))*100;
    EQM(k)=(PrevisaoX-Zteste(k,1))^2;

for i=1:1000
    resb=(bootrsp(Res,1));
    combin=Pcomb+resb;
    Beta=(inv(X'*X))*X'*combin;
    L00=X*Beta; %amostra de pseudo-dados (somado com vetor bootstrap de
resíduos)
    h=[1 c+k]; %h para 1º periodo a frente;
    Ylinear=h*Beta;
    L22=2*Z'-L00;
    PrevBROW=BROWN(L22');
    L050=v*Z'+(1-v)*L00;
    PrevAES=AES(L050');
    PrevBROWN=PrevBROW(k,1);
    Asim=AEScoluna(L22');
    Asimples=Asim(1:c,1);
    Abro=BROWNcoluna(L050');
    Abrow=Abro(1:c,1);
    Pcomb=(L00+Asimples+Abrow)/3;
    PrevTheta=(PrevAES+Ylinear+PrevBROWN)/3;
    Yy1(i)=PrevTheta;
    EQM1(i)=(PrevTheta-PrevisaoX)^2;

end
EST1=full(Yy1)'; %estimativas para (RL+AES)/2 no 1º periodo;
dpadiaoBOOT(k)=std(EST1);
EQMboot(k)=mean(full(EQM1));

alpa=0.05;
INT=sort(EST1'); %ordenação dos valores da amostra bootstrap em ordem crescente;
q1=1000*alpa/2;
q2=1000-q1+1;
lo1=INT(1,q1);
up1=INT(1,q2);

AL(k)=PAES;
YLin(k)=YlinearP;
AB(k)=PBROW;
inf(k)=lo1;
sup(k)=up1;
pthet(k)=PrevisaoX;

```

```
end
```

```
LTheta0=full(YLin)';
LTheta05=full(AB)';
LTheta2=full(AL)';
Previsao=full(pthet)';
intinf=full(inf)';
intsup=full(sup)';
MAPES=full(MAPE)';
EQMs=full(EQM)';
stPADRAO=full(dpadraoBOOT)';
EQMbootstrap=full(EQMboot)';
```

```
disp(' período L(Theta=0) L(Theta=v) L(Theta=2) Previsao inf sup')
```

```
[(1:n)' LTheta0 LTheta05 LTheta2 Previsao intinf intsup ]
```

```
disp(' MAPE EQM STD EQMboot')
[MAPES EQMs stPADRAO EQMbootstrap]
MAPEmedio=mean(MAPES)
EQMmedio=mean(EQMs)
EQMbootmedio=mean(EQMbootstrap)
```

```
%GRAFICOS;
```

```
xl=[1:c];
xll=[1:c+n];
xp=[c+1:c+n];
```

```
plot(xl,Z), legend('S.T.'), title('Série Temporal')
```

```
plot(xl,Z,'b-',xll,[L0;LTheta0]','k--',xll,[L05;LTheta05]','g-',xll,[L2;LTheta2]','k-',xp,Previsao','r:')
legend('S.T.','L(\Theta=0)','L(\Theta=0.5)','L(\Theta=2)','PREVISÃO')
title('SÉRIE TEMPORAL E PREVISÃO THETA'),Xlabel('MESES'),Ylabel('OBSERVAÇÃO')
%%
```

```
if k==1
    figure
    histfit(EST1)
    box off
    xlabel('PREVISÕES THETA ESTIMADAS') % legenda do eixo horizontal
    ylabel('FREQUÊNCIA') %legenda do eixo vertical
    title('DISTRIBUIÇÃO BOOTSTRAP')
end
```

```
end
```

```
end
```

```
save
diary("")
```

APÊNDICE B - PROGRAMA PARA A PREVISÃO DE SÉRIE TEMPORAL ATRAVÉS DO MÉTODO DE ALISAMENTO EXPONENCIAL SIMPLES UTILIZANDO O SOFTWARE MATLAB

```

function[out]=AES(Z)
Z=Z';
[r,c]=size(Z);
for k=1:90
    clear AES
    alphas=[0.1:0.01:0.999]';
    alpha=alphas(k,1);

    for i=1:r;
        for j=1:c;
            AES(1,1)=Z(1,1); %primeiro
            AES(i+1,j)=alpha*Z(i,j)+(1-alpha)*AES(i,j);
            AES;
        end
    end
    Erro=abs(Z-AES(1:r,1));
    menor_erro=sum(Erro);
    if k==1
        melhor_erro=menor_erro;
        melhor_alpha=alpha;
    else
        if menor_erro < melhor_erro
            melhor_erro=menor_erro;
            melhor_alpha=alpha;
        end;
    end;
end
for i=1:r;
    for j=1:c;
        AES(1,1)=Z(1,1); %primeiro valor
        AES(i+1,j)=melhor_alpha*Z(i,j)+(1-melhor_alpha)*AES(i,j);
        AES;
    end
end
AISim=full(AES);
out=AISim(r+1,1);

```

APÊNDICE C - PROGRAMA PARA A PREVISÃO DE SÉRIE TEMPORAL ATRAVÉS DO MÉTODO DE ALISAMENTO EXPONENCIAL DE BROWN UTILIZANDO O SOFTWARE MATLAB

```

function[out]=BROWN(Z)

Z=Z';
[r,c]=size(Z);
for k=1:99
    clear AES
    alphas=[0.01:0.01:0.99]';
    alpha=alphaa(k,1);
    for i=1:r;
        for j=1:c;
            AES(1,1)=Z(1,1); %primeiro
            AES(i+1,j)=alpha*Z(i,j)+(1-alpha)*AES(i,j);
            AES;
        end
    end
    Erro=abs(Z-AES(1:r,1));
    menor_erro=sum(Erro);
    if k==1
        melhor_erro=menor_erro;
        melhor_alpha=alpha;
    else
        if menor_erro < melhor_erro
            melhor_erro=menor_erro;
            melhor_alpha=alpha;
        end;
    end;
end
for i=1:r;
    for j=1:c;
        AES(1,1)=Z(1,1); %primeiro valor
        AES(i+1,j)=melhor_alpha*Z(i,j)+(1-melhor_alpha)*AES(i,j);
        AES;
    end
end
AISim=full(AES);
Alexsimples=AISim(2:r+1,j);
melhor_alpha;
for i=1:r
    for j=1:c

Zbrown(1,1)=Alexsimples(1,1);
Zbrown(i+1,j)=melhor_alpha*Alexsimples(i,j)+(1-melhor_alpha)*Zbrown(i,j);

        Zbrown;

    end
end

```

```

Zbrownn=Zbrown(2:r+1,j);

for i=1:r
    for j=1:c

        at(i,j)=2*Alexsimples(i,j)-Zbrownn(i,j);
        bt(i,j)=(0.2/(1-0.2))*(Alexsimples(i,j)-Zbrownn(i,j));

    end
end
a=at(1:i,j);
b=bt(1:i,j);

for h=1:24
    PBROWN(h)=a(i)+h*b(i);
end
PBROWNN=full(PBROWN');
out=PBROWNN;

```

APÊNDICE D - PROGRAMA PARA A PREVISÃO DE SÉRIE TEMPORAL ATRAVÉS DO MÉTODO DE ALISAMENTO EXPONENCIAL DE HOLT UTILIZANDO O SOFTWARE MATLAB

```

function[out]=HOLT(Z)
Z=Z';
[r,q]=size(Z);
St(1,1)=Z(1,1);
bt(1,1)=((Z(2,1)-Z(1,1))/2)+((Z(4,1)-Z(3,1))/2);
for k=1:99
    clear HOLT
    alphas=[0.01:0.01:0.99]';
    A=alpha(k,1);
    for m=1:99
        beta=[0.01:0.01:0.99]';
        C=beta(m,1);
        for i=1:r-1
            for j=1:q

                St(i+1,j)=A*Z(i+1,j)+(1-A)*(St(i,j)+bt(i,j));
                bt(i+1,j)=C*(St(i+1,j)-St(i,j))+(1-C)*bt(i,j);
                previsaoZ(i+1,j)=St(i,j)+bt(i,j);
                previsaoZ;
                HOLTPRE=previsaoZ(2:i+1,j);

            end
        end
        HOLTPREV=[Z(1,1);HOLTPRE];
        Erro=abs(Z-HOLTPREV);
        menor_erro=sum(Erro);
        if k==1
            melhor_erro=menor_erro;
            melhor_A=A;
            melhor_C=C;
        else
            if menor_erro < melhor_erro
                melhor_erro=menor_erro;
                melhor_A=A;
                melhor_C=C;
            end;
        end;
    end
end
end

for i=1:r-1;
    for j=1:q

        St(i+1,j)=A*Z(i+1,j)+(1-A)*(St(i,j)+bt(i,j));
        bt(i+1,j)=C*(St(i+1,j)-St(i,j))+(1-C)*bt(i,j);
        St;
    end
end

```



```
    bt;  
  
    end  
end  
Stt=St(1:i+1,j);  
btt=bt(1:i+1,j);  
  
Sttt=Stt(i+1,j);  
bttt=btt(i+1,j);  
  
for h=1:24  
    HOLT(h)=Sttt+bttt*h;  
  
end  
HOLTT=full(HOLT');  
format short  
out=HOLTT;
```

APÊNDICE E - PROGRAMA PARA A OTIMIZAÇÃO DO COEFICIENTE THETA UTILIZANDO O SOFTWARE MATLAB

```

function[out]=Ltheta3(Z)
Z=Z';
[r,c]=size(Z);
x=1:r;
X=[(ones(r,1)) x'];      %matriz do modelo de ordem nxp;
Beta=(inv(X'*X))*X'*Z;    %obtenção do parâmetro B;
L0=X*Beta;
L2=2*Z-L0;

AESc=AEScoluna(L2');
AEScc=AESc(1:r,1);

for k=1:18
    clear combi
    T=[0.1;0.2;0.3;0.4;0.5;0.6;0.7;0.8;0.9;1.1;1.2;1.3;1.4;1.5;1.6;1.7;1.8;1.9];
    theta=T(k,1);

    for i=1:r;
        for j=1:c;

            LTheta=theta*Z+(1-theta)*AEScc;
            LTB=BROWNcoluna(LTheta');
            LTBb=LTB(1:r,1);
            combi(i,j)=(L0(i,j)+AEScc(i,j)+LTBb(i,j))/3;
            combi;
        end
    end

    Erro=abs(Z-combi);
    menor_erro=sum(Erro);
    if k==1
        melhor_erro=menor_erro;
        melhor_theta=theta;
    else
        if menor_erro < melhor_erro
            melhor_erro=menor_erro;
            melhor_theta=theta;
        end;
    end;
end
er=sqrt(menor_erro)
SAIDA=melhor_theta

% ultima linha BROWN;

BROWc=BROWNcoluna(L2');
BROWcc=BROWc(1:r,1);

```

```

for k=1:18
    clear combi
    T1=[0.1;0.2;0.3;0.4;0.5;0.6;0.7;0.8;0.9;1.1;1.2;1.3;1.4;1.5;1.6;1.7;1.8;1.9];
    theta1=T1(k,1);

    for i=1:r;
        for j=1:c;

            LTheta=theta*Z+(1-theta)*BROWcc;
            LTB=AEScoluna(LTheta');
            LTBB=LTB(1:r,1);
            combi1(i,j)=(L0(i,j)+BROWcc(i,j)+LTBB(i,j))/3;
            combi1;
        end
    end

    Erro1=abs(Z-combi1);
    menor_erro1=sum(Erro1);
    if k==1
        melhor_erro1=menor_erro1;
        melhor_theta1=theta1;
    else
        if menor_erro1 < melhor_erro1
            melhor_erro1=menor_erro1;
            melhor_theta1=theta1;
        end;
    end;
end
er1=sqrt(menor_erro1)
SAIDA1=melhor_theta1

```

APÊNDICE F - PROGRAMA GERADOR DE AMOSTRAS POR REAMOSTRAGEM COM REPOSIÇÃO

```
function[out]=bootrsp(entrada,B)

%função[saída]=bootrsp(entrada, B)

if (existe('B')~=1), B=1;
end;
if (existe('entrada')~=1), erro('Fornecer dados de Entrada');
end;

s=size(in);
if length(s)>2,
    erro('Entrada de dados pode ser somente vetor ou matriz ');
end;
if min(s)==1,
    out=in(ceil(max(s)*rand(max(s),B)));
else
    out=in(ceil(s(1)*s(2)*rand(s(1),s(2),B)));
end;
```